Journal of Medical Internet Research

Journal Impact Factor (JIF) (2022): 7.4
Volume 21 (2019), Issue 10 ISSN 1438-8871 Editor in Chief: Gunther Eysenbach, MD, MPH

Contents

Viewpoints

Beyond the Impact Factor: Reflecting on Twenty Years of Leading Efforts in Research, Innovation in Publishing, and Investment in People (e16390)
John Torous. . . . 7

Increasing the Impact of JMIR Journals in the Attention Economy (e16172)
Ricky Leung. . . . 10

The Role of the Sharing Economy and Artificial Intelligence in Health Care: Opportunities and Challenges (e13469)
Huailiang Wu, Nga-Kwo Chan, Casper Zhang, Wai-Kit Ming. . . . 64

Trust Me, I'm a Chatbot: How Artificial Intelligence in Health Care Fails the Turing Test (e16222)
John Powell. . . . 708

Role of the Internet in Solving the Last Mile Problem in Medicine (e16385)
Bradford Hesse. . . . 712

The Importance of Health Information on the Internet: How It Saved My Life and How it Can Save Yours (e16690)
Andre Kushniruk. . . . 718

A Research Roadmap: Connected Health as an Enabler of Cancer Patient Support (e14360)
Gabriel Signorelli, Fedor Lehocki, Matilde Mora Fernández, Gillian O'Neill, Dominic O'Connor, Louise Brennan, Francisco Monteiro-Guerra, Alejandro Rivero-Rodriguez, Santiago Hors-Fraile, Juan Munoz-Penas, Mercè Bonjorn Dalmau, Jorge Mota, Ricardo Oliveira, Bela Mrinakova, Silvia Putekova, Naiara Muro, Francisco Zambrana, Juan Garcia-Gomez. . . . 783

Reviews

Rehabilitation, the Great Absentee of Virtual Coaching in Medical Care: Scoping Review (e12805)
Peppino Tropea, Hannes Schlieter, Irma Sterpi, Elda Judica, Kai Gand, Massimo Caprino, Inigo Gabilondo, Juan Gomez-Esteban, Stefan Busnatu, Crina Sinescu, Sofoklis Kyriazakos, Sadia Anwar, Massimo Corbo. . . . 14

Using Patient and Family Engagement Strategies to Improve Outcomes of Health Information Technology Initiatives: Scoping Review (e14683)
Kevin Leung, Drew Lu-McLean, Craig Kuziemsky, Richard Booth, Sarah Collins Rossetti, Elizabeth Borycki, Gillian Strudwick. . . . 36

Barriers to the Use of Mobile Health in Improving Health Outcomes in Developing Countries: Systematic Review (e13263)
Clemens Kruse, Jose Betancourt, Stephanie Ortiz, Susana Valdes Luna, Inderdeep Bamrah, Narce Segovia. 51

Identifying Frameworks for Validation and Monitoring of Consensual Behavioral Intervention Technologies: Narrative Review (e13606)
François Carbonnel, Gregory Ninot. 171

Methodological Challenges in Randomized Controlled Trials on Smartphone-Based Treatment in Psychiatry: Systematic Review (e15362)
Morten Tønning, Lars Kessing, Jakob Bardram, Maria Faurholt-Jepsen. 254

Electronic Patient-Generated Health Data to Facilitate Disease Prevention and Health Promotion: Scoping Review (e13320)
Vasileios Nittas, Penny Lun, Frederic Ehrler, Milo Puhan, Margot Mütsch. 298

Reaching People With Disabilities in Underserved Areas Through Digital Interventions: Systematic Review (e12981)
Leming Zhou, Bambang Parmanto. 352

Effects of E-Learning in a Continuing Education Context on Nursing Care: Systematic Review of Systematic Qualitative, Quantitative, and Mixed-Studies Reviews (e15118)
Geneviève Rouleau, Marie-Pierre Gagnon, José Côté, Julie Payne-Gagnon, Emilie Hudson, Carl-Ardy Dubois, Julien Bouix-Picasso. 448

Original Papers

Diminishing Effects After Recurrent Use of Self-Guided Internet-Based Interventions in Depression: Randomized Controlled Trial (e14240)
Lara Bücker, Patricia Schnakenberg, Eirini Karyotaki, Steffen Moritz, Stefan Westermann. 68

Promoting Social Connection and Deepening Relations Among Older Adults: Design and Qualitative Evaluation of Media Parcels (e14112)
Isabela Zaine, David Frohlich, Kamila Rodrigues, Bruna Cunha, Alex Orlando, Leonardo Scalco, Maria Pimentel. 85

Development of the Therapeutic Alliance and its Association With Internet-Based Mindfulness-Based Cognitive Therapy for Distressed Cancer Patients: Secondary Analysis of a Multicenter Randomized Controlled Trial (e14065)
Else Bisseling, Linda Cillessen, Philip Spinhoven, Melanie Schellekens, Félix Compen, Marije van der Lee, Anne Speckens. 103

Internet-Based Cognitive Behavioral Therapy for Chronic Fatigue Syndrome Integrated in Routine Clinical Care: Implementation Study (e14037)
Margreet Worm-Smeitink, Arno van Dam, Saskia van Es, Rosalie van der Vaart, Andrea Evers, Michel Wensing, Hans Knoop. 115

Estimating the Impact of Novel Digital Therapeutics in Type 2 Diabetes and Hypertension: Health Economic Analysis (e15814)
Robert Nordyke, Kevin Appelbaum, Mark Berman. 129

Process Evaluation of Nurse-Led Online Self-Management Support for Family Caregivers to Deal With Behavior Changes of a Relative With Dementia (Part 1): Mixed Methods Study (e13002)
Judith Huis in het Veld, Iris van Asch, Bernadette Willemse, Paul-Jeroen Verkade, Anne Pot, Marco Blom, Rob Groot Zwaaftink, Anneke Francke. 140

Development and Efficacy of an Electronic, Culturally Adapted Lifestyle Counseling Tool for Improving Diabetes-Related Dietary Knowledge: Randomized Controlled Trial Among Ethnic Minority Adults With Type 2 Diabetes Mellitus (e13674)
Kathleen Abu-Saad, Havi Murad, Rivka Barid, Liraz Olmer, Arnona Ziv, Nuha Younis-Zeidan, Vered Kaufman-Shriqui, Michal Gillon-Keren, Shmuel Rigler, Yakir Berchenko, Ofra Kalter-Leibovici. . . . 152

Feasibility, Acceptability, and Preliminary Impacts of Web-Based Patient Education on Patients With Schizophrenia Spectrum Disorder: Quasi-Experimental Cluster Study (e13073)
Anna Laine, Maritta Välimäki, Virve Pekurinen, Eliisa Löyttyniemi, Mauri Marttunen, Minna Anttila. . . . 186

A Comparison of Users and Nonusers of a Web-Based Intervention for Carers of Older Persons With Alzheimer Disease and Related Dementias: Mixed Methods Secondary Analysis (e14254)
Wendy Duggleby, Jenny Ploeg, Carrie McAiney, Kathryn Fisher, Kathya Jovel Ruiz, Sunita Ghosh, Shelley Peacock, Maureen Markle-Reid, Allison Williams, Jean Triscott, Jennifer Swindle. . . . 204

Reduced Hospitalizations, Emergency Room Visits, and Costs Associated with a Web-Based Health Literacy, Aligned-Incentive Intervention: Mixed Methods Study (e14772)
Jeffrey Greene, Jolie Haun, Dustin French, Susan Chambers, Robert Roswell. . . . 217

Electronic Health Record–Based Strategy to Promote Medication Adherence Among Patients With Diabetes: Longitudinal Observational Study (e13499)
Stacy Bailey, Amisha Wallia, Sarah Wright, Guisselle Wismer, Alexandra Infanzon, Laura Curtis, Samantha Brokenshire, Arlene Chung, Daniel Reuland, Allison Hahr, Kenneth Hornbuckle, Karen Lockwood, Lori Hall, Michael Wolf. . . . 230

Technology Acceptance and Information System Success of a Mobile Electronic Platform for Nonphysician Clinical Students in Zambia: Prospective, Nonrandomized Intervention Study (e14748)
Sandra Barteit, Florian Neuhann, Till Bärnighausen, Annel Bowa, Sigrid Wolter, Hinson Siabwanta, Albrecht Jahn. . . . 239

Trustworthy Health-Related Tweets on Social Media in Saudi Arabia: Tweet Metadata Analysis (e14731)
Yahya Albalawi, Nikola Nikolov, Jim Buckley. . . . 276

A Web-Based Intervention for Social Media Addiction Disorder Management in Higher Education: Quantitative Survey Study (e14834)
Huseyin Dogan, Helmi Norman, Amen Alrobai, Nan Jiang, Norazah Nordin, Anita Adnan. . . . 317

Improving the Course of Depressive Symptoms After Inpatient Psychotherapy Using Adjunct Web-Based Self-Help: Follow-Up Results of a Randomized Controlled Trial (e13655)
Rüdiger Zwerenz, Carlotta Baumgarten, Jan Becker, Ana Tibubos, Martin Siepmann, Rudolf Knickenberg, Manfred Beutel. . . . 329

Utilization, Safety, and Technical Performance of a Telemedicine System for Prehospital Emergency Care: Observational Study (e14907)
Marc Felzen, Stefan Beckers, Felix Kork, Frederik Hirsch, Sebastian Bergrath, Anja Sommer, Jörg Brokmann, Michael Czaplik, Rolf Rossaint.
3 4 3

Participant-Centered Online Active Surveillance for Adverse Events Following Vaccination in a Large Clinical Trial: Feasibility and Usability Study (e14791)
Sally-Anne Munnoch, Patrick Cashman, Roseanne Peel, John Attia, Alexis Hure, David Durrheim. . . . 367

An Online Survey for Pharmacoepidemiological Investigation (Survey of Non-Medical Use of Prescription Drugs Program): Validation Study (e15830)
Joshua Black, Karilynn Rockhill, Alyssa Forber, Elise Amioka, K May, Colleen Haynes, Nabarun Dasgupta, Richard Dart. . . . 379

Characterizing the Rural Opioid Use Environment in Kentucky Using Google Earth: Virtual Audit (e14923)
Natalie Crawford, Regine Haardöerfer, Hannah Cooper, Izraelle McKinnon, Carla Jones-Harrell, April Ballard, Sierra von Hellens, April Young.
3 9 4

Tailored Web-Based Information for Younger and Older Patients with Cancer: Randomized Controlled Trial of a Preparatory Educational Intervention on Patient Outcomes (e14407)
Minh Nguyen, Ellen Smets, Nadine Bol, Eugène Loos, Hanneke van Laarhoven, Debby Geijsen, Mark van Berge Henegouwen, Kristien Tytgat, Julia van Weert. . . . 403

Difference Between Users and Nonusers of a Patient Portal in Health Behaviors and Outcomes: Retrospective Cohort Study (e13146)
Jing Huang, Yong Chen, J Landis, Kevin Mahoney. . . . 427

The Internet Intervention Patient Adherence Scale for Guided Internet-Delivered Behavioral Interventions: Development and Psychometric Evaluation (e13602)
Fabian Lenhard, Kajsa Mitsell, Maral Jolstedt, Sarah Vigerland, Tove Wahlund, Martina Nord, Johan Bjureberg, Hanna Sahlin, Per Andrén, Kristina Aspvall, Karin Melin, David Mataix-Cols, Eva Serlachius, Jens Högström. . . . 437

Codifying Online Social Support for Breast Cancer Patients: Retrospective Qualitative Assessment (e12880)
Jude Mikal, Stuart Grande, Michael Beckstrand. . . . 467

Difficulties Encountered by People With Depression and Anxiety on the Web: Qualitative Study and Web-Based Expert Survey (e12514)
Renaldo Bernard, Carla Sabariego, Alarcos Cieza. . . . 484

Patient Engagement in Medical Research Among Older Adults: Analysis of the Health Information National Trends Survey (e15035)
Lynette Gerido, Xiang Tang, Brittany Ernst, Aisha Langford, Zhe He. . . . 503

Public Concern About Monitoring Twitter Users and Their Conversations to Recruit for Clinical Trials: Survey Study (e15455)
Katja Reuter, Yifan Zhu, Praveen Angyan, NamQuyen Le, Akil Merchant, Michael Zimmer. . . . 514

Developing a Hypothetical Implementation Framework of Expectations for Monitoring Early Signs of Psychosis Relapse Using a Mobile App: Qualitative Study (e14366)
Stephanie Allan, Simon Bradstreet, Hamish Mcleod, John Farhall, Maria Lambrou, John Gleeson, Andrea Clark, EMPOWER Group, Andrew Gumley. . . . 530

Antismoking Advertisements and Price Promotions and Their Association With the Urge to Smoke and Purchases in a Virtual Convenience Store: Randomized Experiment (e14143)
Lauren Dutra, James Nonnemaker, Brian Bradfield, Nathaniel Taylor, Jamie Guillory, Ashley Feld, Annice Kim. . . . 544

A Virtual Counseling Application Using Artificial Intelligence for Communication Skills Training in Nursing Education: Development Study (e14658)
Shefaly Shorey, Emily Ang, John Yap, Esperanza Ng, Siew Lau, Chee Chui. . . . 556

The Psychometric Properties of the Chinese eHealth Literacy Scale (C-eHEALS) in a Chinese Rural Population: Cross-Sectional Validation Study (e15720)
Zhihao Ma, Mei Wu. . . . 570

Results of MyPlan 2.0 on Physical Activity in Older Belgian Adults: Randomized Controlled Trial (e13219)
Delfien Van Dyck, Karel Herman, Louise Poppe, Geert Crombez, Ilse De Bourdeaudhuij, Freja Gheysen. . . . 584

Cost and Effectiveness of Blended Versus Standard Cognitive Behavioral Therapy for Outpatients With Depression in Routine Specialized Mental Health Care: Pilot Randomized Controlled Trial (e14261)
Lisa Kooistra, Jenneke Wiersma, Jeroen Ruwaard, Koen Neijenhuijs, Joran Lokkerbol, Patricia van Oppen, Filip Smit, Heleen Riper. . . . 597

Growing Disparities in Patient-Provider Messaging: Trend Analysis Before and After Supportive Policy (e14976)
Nicole Senft, Evan Butler, Jordan Everson. . . . 614

Psychosocial Factors Affecting Artificial Intelligence Adoption in Health Care in China: Cross-Sectional Study (e14316)
Tiantian Ye, Jiaolong Xue, Mingguang He, Jing Gu, Haotian Lin, Bin Xu, Yu Cheng. . . . 626

Use of a Web-Based Dietary Assessment Tool (RiksmatenFlex) in Swedish Adolescents: Comparison and Validation Study (e12572)
Anna Lindroos, Jessica Petrelius Sipinen, Cecilia Axelsson, Gisela Nyberg, Rikard Landberg, Per Leanderson, Marianne Arnemo, Eva Warensjö Lemming. . . . 648

Identifying the Most Autonomy-Supportive Message Frame in Digital Health Communication: A 2x2 Between-Subjects Experiment (e14074)
Eline Smit, Chamoetal Zeidler, Ken Resnicow, Hein de Vries. . . . 664

Home Virtual Visits for Outpatient Follow-Up Stroke Care: Cross-Sectional Study (e13734)
Ramana Appireddy, Sana Khan, Chad Leaver, Cally Martin, Albert Jin, Bryce Durafourt, Stephen Archer. . . . 676

Hospital-Based Back Surgery: Geospatial-Temporal, Explanatory, and Predictive Models (e14609)
Lawrence Fulton, Clemens Kruse. . . . 687

Towards a Stakeholder-Oriented Blockchain-Based Architecture for Electronic Health Records: Design Science Research Study (e13585)
Jan Beinke, Christian Fitte, Frank Teuteberg. . . . 754

The Effect of Robot Attentional Behaviors on User Perceptions and Behaviors in a Simulated Health Care Interaction: Randomized Controlled Trial (e13667)
Deborah Johanson, Ho Ahn, Bruce MacDonald, Byeong Ahn, JongYoon Lim, Euijun Hwang, Craig Sutherland, Elizabeth Broadbent. . . . 798

A Protocol-Driven, Bedside Digital Conversational Agent to Support Nurse Teams and Mitigate Risks of Hospitalization in Older Adults: Case Control Pre-Post Study (e13440)
Nicholas Bott, Sharon Wexler, Lin Drury, Chava Pollak, Victor Wang, Kathleen Scher, Sharon Narducci. . . . 813

White Paper

An Integrated Brain-Machine Interface Platform With Thousands of Channels (e16194)
Elon Musk, Neuralink. . . . 722

Commentaries

The Ethical and Responsible Development and Application of Advanced Brain Machine Interfaces (e16321)
Andrew Maynard, Marissa Scragg. . . . 736

The Reconnecting the Hand and Arm with Brain (ReHAB) Commentary on “An Integrated Brain-Machine Interface Platform With Thousands of Channels” (e16339)
Robert Kirsch, A Ajiboye, Jonathan Miller. . . . 744

From Novel Technology to Novel Applications: Comment on “An Integrated Brain-Machine Interface Platform With Thousands of Channels” by Elon Musk and Neuralink (e16356)
Alexander Pisarchik, Vladimir Maksimenko, Alexander Hramov. . . . 747

Tutorial

Building a Secure Biomedical Data Sharing Decentralized App (DApp): Tutorial (e13601)
Matthew Johnson, Michael Jones, Mark Shervey, Joel Dudley, Noah Zimmerman. 768

Corrigenda and Addendas

Correction: Adherence and Satisfaction of Smartphone- and Smartwatch-Based Remote Active Testing and Passive Monitoring in People With Multiple Sclerosis: Nonrandomized Interventional Feasibility Study (e16287)
Luciana Midaglia, Patricia Mulero, Xavier Montalban, Jennifer Graves, Stephen Hauser, Laura Julian, Michael Baker, Jan Schadrack, Christian Gossens, Alf Scotland, Florian Lipsmeier, Johan van Beek, Corrado Bernasconi, Shibeshih Belachew, Michael Lindemann. 825

Addendum to the Acknowledgements: Mood Prediction of Patients With Mood Disorders by Machine Learning Using Passive Digital Phenotypes Based on the Circadian Rhythm: Prospective Observational Cohort Study (e15966)
Chul-Hyun Cho, Taek Lee, Min-Gwan Kim, Hoh In, Leen Kim, Heon-Jeong Lee. 827

Viewpoint

Beyond the Impact Factor: Reflecting on Twenty Years of Leading Efforts in Research, Innovation in Publishing, and Investment in People

John Torous[1], MBI, MD

Division of Digital Psychiatry, Department of Psychiatry, Beth Israel Deaconess Medical Center, Boston, MA, United States

Corresponding Author:
John Torous, MBI, MD
Division of Digital Psychiatry, Department of Psychiatry
Beth Israel Deaconess Medical Center
75 Fenwood Rd
Boston, MA, 02115
United States
Phone: 1 682 7822
Email: jtorous@bidmc.harvard.edu

Abstract

This viewpoint celebrates the accomplishments of the Journal of Medical Internet Research (JMIR) on its twentieth anniversary and reviews accomplishments around research publications, journal innovation, and supporting people.

(J Med Internet Res 2019;21(10):e16390) doi:10.2196/16390

KEYWORDS

JMIR; publishing; eHealth; digital health; digital medicine; knowledge dissemination

“Form follows function” is an aphorism that has transcended architecture, computer science, and now even digital health. Reflecting on the twentieth anniversary of the Journal of Medical Internet Research (JMIR), “form follows function” also well explains the success and impact of the journal in promoting high quality research, supporting innovation in publication, and supporting a generation of young investigators.

In the 20 years JMIR has represented the field of digital health, the nature of digital health research has evolved and transformed. Today’s papers often focus on machine learning, virtual reality, smartphone apps, and other topics that were nonexistent or nascent in health care research in 1999 when the journal began, or in 2001 when the journal outlined a clear definition for electronic health (eHealth) [1]. Yet today, the journal continues to function as a hub for digital health research. The prescient nature of JMIR in publishing not only high quality but also innovative articles is exemplified by a sample of papers from 20 years prior, with topics including evaluation of consumer health tools [2], neural nets for medical decision making [3], and online prescribing [4], all of which were published in 1999. With its reputation for high quality publications, the journal has offered clinicians, patients, researchers, entrepreneurs, and policy makers a window into best practices and current evidence for digital health. While it is impossible to know what the critical topics and forms of research in digital health will be in the future, it is possible to state that JMIR will cover those with the same high standards and functions it has stood by for its first 20 years.

However, JMIR is more than high quality research. The interdisciplinary nature, rapid paradigm shifts, and global impact of digital health requires a journal to adopt a unique form to meet its function in this space. For 20 years, JMIR has been rethinking how a journal can support the digital mental health community with leading efforts in open access publication, preprints, an easy to access Web and mobile layout, social media outreach and dissemination, crowdfunding, and offering peer reviewers publication credits, among others. The journal’s adaptable online format can support research from not only the health fields but also engineering, computer sciences, design, implementation, policy, and other unique disciplines. Its unique functioning has created a form that offers the ability to publish unique special issues [5] on topics that otherwise would likely not find a voice. The format of JMIR has adapted with changes in the way people access, share, and read research papers. New innovations such as the new JMIRx journal series ("superjournals" or overlay journals for Preprint servers, announced in this 20th Anniversary Theme Issue), highlight ongoing efforts to expand the form of the journal towards an open science platform.

Even beyond high quality research and innovative publishing models, JMIR's legacy of 20 years includes supporting people. The journal has forged a reputation for offering patients a voice in the medical world [6,7]. It also serves as a forum to engage all members of the digital health ecosystem to share their research, perspectives, and ideas, whether as a world-leading research team [8] or simply as a person with a thoughtful and important perspective to share. On a personal note, the journal has supported my own career in digital health, offering a venue to share my early research in 2014 [9] and later to serve on the editorial board, before then becoming editor-in-chief of JMIR Mental Health. I can also count many other junior investigators for whom JMIR has served as a mentor and adapted its form to help support their interests and exploration of digital health. This commitment towards supporting people of all natures and backgrounds may be the most enduring and impressive aspect of the first 20 years of JMIR.

Reflecting on the twentieth anniversary of JMIR, it is clear that "form follows function" only begins to describe the innovation and impact of the journal. From leading efforts in research, innovation in publication, and investment in people, the field has grown because of JMIR and its support for digital health. While the format of JMIR will continue to adapt to lead the field, its function as a leading force will remain constant and it will continue to act as a beacon of certainty in the always evolving world of digital health.

Conflicts of Interest

None declared.

References

1. Eysenbach G. What is e-health? J Med Internet Res 2001 Jun 18;3(2):E20 [FREE Full text] [doi: 10.2196/jmir.3.2.e20] [Medline: 11720962]
2. Cui L. Rating health web sites using the principles of citation analysis: a bibliometric approach. J Med Internet Res 1999 Aug 5;1(1):E4 [FREE Full text] [doi: 10.2196/jmir.1.1.e4] [Medline: 11720913]
3. Kanagaratnam B, Lavelle S, Comerford R. FTO1/362: Using Neural Nets in Medical Decision Making. J Med Internet Res 1999 Sep 19;1:e26. [doi: 10.2196/jmir.1.suppl1.e26]
4. Eysenbach G. Online prescribing of sildanefil (Viagra) on the world wide web. J Med Internet Res 1999 Dec 31;1(2):E10 [FREE Full text] [doi: 10.2196/jmir.1.2.e10] [Medline: 11720919]
5. Calvo RA, Dinakar K, Picard R, Christensen H, Torous J. Toward Impactful Collaborations on Computing and Mental Health. J Med Internet Res 2018 Feb 09;20(2):e49 [FREE Full text] [doi: 10.2196/jmir.9021] [Medline: 29426812]
6. Torous J, Roux S. Patient-Driven Innovation for Mobile Mental Health Technology: Case Report of Symptom Tracking in Schizophrenia. JMIR Ment Health 2017 Jul 06;4(3):e27 [FREE Full text] [doi: 10.2196/mental.7911] [Medline: 28684386]
7. Chiauzzi E, Newell A. Mental Health Apps in Psychiatric Treatment: A Patient Perspective on Real World Technology Usage. JMIR Ment Health 2019 Apr 22;6(4):e12292 [FREE Full text] [doi: 10.2196/12292] [Medline: 31008711]
8. Bosl W, Mandel J, Jonikas M, Ramoni RB, Kohane IS, Mandl KD. Scalable decision support at the point of care: a substitutable electronic health record app for monitoring medication adherence. Interact J Med Res 2013 Jul 22;2(2):e13 [FREE Full text] [doi: 10.2196/ijmr.2480] [Medline: 23876796]
9. Torous J, Friedman R, Keshavan M. Smartphone ownership and interest in mobile applications to monitor symptoms of mental health conditions. JMIR Mhealth Uhealth 2014 Jan 21;2(1):e2 [FREE Full text] [doi: 10.2196/mhealth.2994] [Medline: 25098314]

Abbreviations

eHealth: electronic health
JMIR: Journal of Medical Internet Research

Edited by G Eysenbach; submitted 24.09.19; peer-reviewed by C Khoury; comments to author 13.10.19; revised version received 14.10.19; accepted 17.10.19; published 31.10.19.

Please cite as:
Torous J
Beyond the Impact Factor: Reflecting on Twenty Years of Leading Efforts in Research, Innovation in Publishing, and Investment in People
J Med Internet Res 2019;21(10):e16390
URL: http://www.jmir.org/2019/10/e16390/
doi:10.2196/16390
PMID:31674922

Viewpoint

Increasing the Impact of JMIR Journals in the Attention Economy

Ricky Leung[1], PhD

University at Albany, School of Public Health, Rensselaer, NY, United States

Corresponding Author:
Ricky Leung, PhD
University at Albany
School of Public Health
1 University Pl, Room 181
Rensselaer, NY, 12144
United States
Phone: 1 6083348781
Email: rleung122@hotmail.com

Abstract

The Journal of Medical Internet Research (JMIR) has attained remarkable achievements in the past twenty years. By depth, JMIR has published the most impactful research in medical informatics and is top ranked in the field. By width, JMIR has spun off to about thirty sister journals to cover topics such as serious games, mobile health, public health, surveillance, and other medical areas. With ever-increasing data and research findings, academic publishers need to be competitive to win readers' attention. While JMIR is well-positioned in the field, the journal will need more creative strategies to increase its attention base and maintain its leading position. Viable strategies include the creation of online collaborative spaces, the engagement of more diverse audience from less traditional channels, and partnerships with other publishers and academic institutes. Doing so could also enable JMIR researchers to turn research insights into practical strategies to improve personal health and medical services.

(J Med Internet Res 2019;21(10):e16172) doi:10.2196/16172

KEYWORDS

JMIR; medical informatics; digital health; publishing; knowledge translation; peer-to-peer community; impact

The Journal of Medical Internet Research (JMIR) has engaged many health researchers since the journal published its inaugural issue in August 1999 [1]. JMIR is now ranked number one in the field of medical informatics, and the JMIR publishing office has issued close to 30 additional sister journals, including JMIR mHealth and uHealth, the Journal of Serious Games, and others. These are remarkable achievements. In this short article, I provide a quick review to highlight the important role of JMIR in promoting innovative research in the field of medical informatics. Building on its achievements, JMIR is well-positioned not only to maintain its leading position in medical informatics but also to contribute excellent research to the academic community at large.

Medical informatics researchers have long recognized that the internet is a big treasure of useful data to improve medical practice and health behaviors. For example, JMIR research published in earlier years studied how practitioners could utilize available statistical data on the internet to measure health quality [2-4]. As online tools became accessible, researchers analyzed whether the internet could serve as an effective platform for primary data collection through surveys [5,6], experiments [7], and even interviews [8]. Some researchers evaluated the efficacy of online medical treatment and found positive results in terms of reducing depression and other medical symptoms [9,10]. This research has paved the way for the emerging subfield of real-world data/evidence [11]. Pharmaceutical companies and public health agencies, such as the US Food and Drug Administration, have expended considerable efforts to promote real-world evidence to supplement and even replace expensive clinical trials.

JMIR research has applied some of the most sophisticated methods in the field. For instance, some researchers examined unstructured data with advanced text-mining techniques [12]. This line of research can detect the sentiments of participants in social media platforms [13] and can determine other deep meanings embedded in qualitative data. In addition, researchers are now able to build robots [14] and apply artificial intelligence [15] to conduct research projects that were not feasible in the past, such as using machine learning algorithms to capture real-time data from social media channels [16]. To reduce the impact of fake news on health outcomes, other JMIR researchers evaluated the veracity of news reports from multiple channels [17].

The rapid development of connected devices, such as wearable technology, smart appliances, and body sensors, has presented new opportunities and challenges for medical informatics research. JMIR has already published some exciting research about the Internet of Things which has multilevel implications for patients and health providers. For instance, many recent studies published in JMIR have shown a strong patient-centered orientation [18]. These studies focused on how the internet has enabled patients to incorporate first-hand experience into research to increase its practical value, whether from actual health care experience, usage of Internet-based devices, or information from peer groups. Funding agencies such as the Patient-Centered Outcomes Research Institute and the National Institutes of Health have paid attention to this research published in JMIR [19].

Beyond personal health, new information technology has improved the quality of health care delivery. For example, interorganizational networks have allowed hospitals to access patients' records from different healthcare settings, facilitating the transition from electronic medical records to electronic health records [20]. Many studies published in JMIR were products of collaborative networks in health care. More recently, research published in JMIR has built ambitious frameworks useful for studying issues at the community and societal levels. Researchers have used these new frameworks to study such important global issues as aging [21], climate change [22], poverty [23] and sustainability [24].

The availability of big data in the medical sector has led to the challenge of limited attention for researchers [25]. While research institutes and established journals have employed aggressive marketing campaigns to attract attention from existing and potential consumers [26], medical researchers now receive enormous amounts of information from social media, personal devices, and other online platforms. In this context, JMIR needs to come up with creative strategies to increase its intellectual breadth and depth to maintain its leading position in the field. One viable strategy is to provide more collaborative space for JMIR subscribers and potential authors. That is, instead of serving only as a publisher of completed research, JMIR could also become a platform to promote collaboration for early-stage research. In fact, JMIR has recently started a Digital Health Community [27] within the JMIR Publications Knowledge Base and Help Center [28] to support the germination of new ideas and allow authors to provide feedback to the journal. The question is how to engage digital health researchers in platforms like that. To encourage increased participation in this community, JMIR may offer Karma points (already implemented to reward reviewers) that can serve as credits to reduce publication fees. Besides that, the JMIR editorial team can reach out and engage researchers from related fields to increase the journal's impact. Increased attendance of academic conferences and professional meetings may be very useful for increasing the exposure of JMIR in various academic communities. Finally, JMIR may consider forming new partnerships with other publishers, professional organizations, academic institutes in different countries, and even funding agencies. Used properly, these strategies can increase the impact of JMIR in the next 20 years - beyond the impact factor.

Conflicts of Interest

None declared.

References

1. Eysenbach G. Welcome to the Journal of Medical Internet Research. J Med Internet Res 1999 Aug 11;1(1):e5 [FREE Full text] [doi: 10.2196/jmir.1.1.e5]
2. Hernández-Borges AA, Macías-Cervi P, Gaspar-Guardado MA, Torres-Alvarez de Arcaya ML, Ruiz-Rabaza A, Jiménez-Sosa A. Can examination of WWW usage statistics and other indirect quality indicators distinguish the relative quality of medical web sites? J Med Internet Res 1999 Aug 11;1(1):e1 [FREE Full text] [doi: 10.2196/jmir.1.1.e1] [Medline: 11720910]
3. Norum J. Evaluation of Norwegian cancer hospitals web sites and explorative survey among cancer patients on their use of the internet. J Med Internet Res 2001 Dec 26;3(4):e30 [FREE Full text] [doi: 10.2196/jmir.3.4.e30] [Medline: 11772545]
4. Jeannot J, Scherer F, Pittet V, Burnand B, Vader J. Use of the World Wide Web to implement clinical practice guidelines: a feasibility study. J Med Internet Res 2003 Jun 13;5(2):e12 [FREE Full text] [doi: 10.2196/jmir.5.2.e12] [Medline: 12857668]
5. Brøgger J, Nystad W, Cappelen I, Bakke P. No increase in response rate by adding a web response option to a postal population survey: a randomized trial. J Med Internet Res 2007 Dec 31;9(5):e40 [FREE Full text] [doi: 10.2196/jmir.9.5.e40] [Medline: 18174120]
6. Eysenbach G, Wyatt J. Using the Internet for surveys and health research. J Med Internet Res 2002 Nov 22;4(2):e13 [FREE Full text] [doi: 10.2196/jmir.4.2.e13] [Medline: 12554560]
7. Overberg R, Otten W, de Man A, Toussaint P, Westenbrink J, Zwetsloot-Schonk B. How breast cancer patients want to search for and retrieve information from stories of other patients on the internet: an online randomized controlled experiment. J Med Internet Res 2010 Mar 09;12(1):e7 [FREE Full text] [doi: 10.2196/jmir.1215] [Medline: 20215101]
8. Powell J, Inglis N, Ronnie J, Large S. The characteristics and motivations of online health information seekers: cross-sectional survey and qualitative interview study. J Med Internet Res 2011 Feb 23;13(1):e20 [FREE Full text] [doi: 10.2196/jmir.1600] [Medline: 21345783]

9. Meyer B, Berger T, Caspar F, Beevers CG, Andersson G, Weiss M. Effectiveness of a novel integrative online treatment for depression (Deprexis): randomized controlled trial. J Med Internet Res 2009 May 11;11(2):e15 [FREE Full text] [doi: 10.2196/jmir.1151] [Medline: 19632969]
10. Rooke S, Copeland J, Norberg M, Hine D, McCambridge J. Effectiveness of a self-guided web-based cannabis treatment program: randomized controlled trial. J Med Internet Res 2013 Feb 15;15(2):e26 [FREE Full text] [doi: 10.2196/jmir.2256] [Medline: 23470329]
11. McCall HC, Helgadottir FD, Menzies RG, Hadjistavropoulos HD, Chen FS. Evaluating a Web-Based Social Anxiety Intervention Among Community Users: Analysis of Real-World Data. J Med Internet Res 2019 Jan 10;21(1):e11566 [FREE Full text] [doi: 10.2196/11566] [Medline: 30632965]
12. Celi LA, Ippolito A, Montgomery RA, Moses C, Stone DJ. Crowdsourcing knowledge discovery and innovations in medicine. J Med Internet Res 2014 Sep 19;16(9):e216 [FREE Full text] [doi: 10.2196/jmir.3761] [Medline: 25239002]
13. Greaves F, Ramirez-Cano D, Millett C, Darzi A, Donaldson L. Use of sentiment analysis for capturing patient experience from free-text comments posted online. J Med Internet Res 2013 Nov 01;15(11):e239 [FREE Full text] [doi: 10.2196/jmir.2721] [Medline: 24184993]
14. Broadbent E, Garrett J, Jepsen N, Li Ogilvie V, Ahn HS, Robinson H, et al. Using Robots at Home to Support Patients With Chronic Obstructive Pulmonary Disease: Pilot Randomized Controlled Trial. J Med Internet Res 2018 Feb 13;20(2):e45 [FREE Full text] [doi: 10.2196/jmir.8640] [Medline: 29439942]
15. Contreras I, Vehi J. Artificial Intelligence for Diabetes Management and Decision Support: Literature Review. J Med Internet Res 2018 May 30;20(5):e10775 [FREE Full text] [doi: 10.2196/10775] [Medline: 29848472]
16. Cole-Lewis H, Varghese A, Sanders A, Schwarz M, Pugatch J, Augustson E. Assessing Electronic Cigarette-Related Tweets for Sentiment and Content Using Supervised Machine Learning. J Med Internet Res 2015 Aug 25;17(8):e208 [FREE Full text] [doi: 10.2196/jmir.4392] [Medline: 26307512]
17. Chen L, Wang X, Peng T. Nature and Diffusion of Gynecologic Cancer-Related Misinformation on Social Media: Analysis of Tweets. J Med Internet Res 2018 Oct 16;20(10):e11515 [FREE Full text] [doi: 10.2196/11515] [Medline: 30327289]
18. Dyson MP, Shave K, Fernandes RM, Scott SD, Hartling L. Outcomes in Child Health: Exploring the Use of Social Media to Engage Parents in Patient-Centered Outcomes Research. J Med Internet Res 2017 Mar 16;19(3):e78 [FREE Full text] [doi: 10.2196/jmir.6655] [Medline: 28302593]
19. Sadasivam RS, Borglund EM, Adams R, Marlin BM, Houston TK. Impact of a Collective Intelligence Tailored Messaging System on Smoking Cessation: The Perspect Randomized Experiment. J Med Internet Res 2016 Nov 08;18(11):e285 [FREE Full text] [doi: 10.2196/jmir.6465] [Medline: 27826134]
20. Rodrigues JJPC, de la Torre I, Fernández G, López-Coronado M. Analysis of the security and privacy requirements of cloud-based electronic health records systems. J Med Internet Res 2013 Aug 21;15(8):e186 [FREE Full text] [doi: 10.2196/jmir.2494] [Medline: 23965254]
21. Peek STM, Wouters EJ, Luijkx KG, Vrijhoef HJ. What it Takes to Successfully Implement Technology for Aging in Place: Focus Groups With Stakeholders. J Med Internet Res 2016 May 03;18(5):e98 [FREE Full text] [doi: 10.2196/jmir.5253] [Medline: 27143097]
22. Tizek L, Schielein M, Rüth M, Ständer S, Pereira M, Eberlein B, et al. Influence of Climate on Google Internet Searches for Pruritus Across 16 German Cities: Retrospective Analysis. J Med Internet Res 2019 Jul 12;21(7):e13739 [FREE Full text] [doi: 10.2196/13739] [Medline: 31301128]
23. Gopalan A, Makelarski JA, Garibay LB, Escamilla V, Merchant RM, Wolfe MB, et al. Health-Specific Information and Communication Technology Use and Its Relationship to Obesity in High-Poverty, Urban Communities: Analysis of a Population-Based Biosocial Survey. J Med Internet Res 2016 Jun 28;18(6):e182 [FREE Full text] [doi: 10.2196/jmir.5741] [Medline: 27352770]
24. Greenhalgh T, Wherton J, Papoutsi C, Lynch J, Hughes G, A'Court C, et al. Beyond Adoption: A New Framework for Theorizing and Evaluating Nonadoption, Abandonment, and Challenges to the Scale-Up, Spread, and Sustainability of Health and Care Technologies. J Med Internet Res 2017 Nov 01;19(11):e367 [FREE Full text] [doi: 10.2196/jmir.8775] [Medline: 29092808]
25. Shaikh AR, Butte AJ, Schully SD, Dalton WS, Khoury MJ, Hesse BW. Collaborative biomedicine in the age of big data: the case of cancer. J Med Internet Res 2014 Apr 07;16(4):e101 [FREE Full text] [doi: 10.2196/jmir.2496] [Medline: 24711045]
26. Cugelman B, Thelwall M, Dawes P. Online interventions for social marketing health behavior change campaigns: a meta-analysis of psychological architectures and adherence factors. J Med Internet Res 2011 Feb 14;13(1):e17 [FREE Full text] [doi: 10.2196/jmir.1367] [Medline: 21320854]
27. JMIR Publications. Digital Health Community. 2019. URL: https://support.jmir.org/hc/en-us/community/topics [accessed 2019-10-24]
28. JMIR Publications. JMIR Publications Knowledge Base and Help Center. 2019. URL: https://support.jmir.org/hc/en-us/ [accessed 2019-10-24]

Abbreviations

JMIR: Journal of Medical Internet Research

Edited by G Eysenbach; submitted 06.09.19; peer-reviewed by P Zhang, J Torous; comments to author 03.10.19; revised version received 14.10.19; accepted 22.10.19; published 31.10.19.

Please cite as:
Leung R
Increasing the Impact of JMIR Journals in the Attention Economy
J Med Internet Res 2019;21(10):e16172
URL: http://www.jmir.org/2019/10/e16172/
doi:10.2196/16172
PMID:31674916

Review

Rehabilitation, the Great Absentee of Virtual Coaching in Medical Care: Scoping Review

Peppino Tropea[1], PhD; Hannes Schlieter[2], PhD; Irma Sterpi[1], PhD; Elda Judica[1], MD; Kai Gand[2], PhD; Massimo Caprino[1], MSc; Inigo Gabilondo[3,4], MD, PhD; Juan Carlos Gomez-Esteban[3], MD; Stefan Busnatu[5], MD, PhD; Crina Sinescu[5], MD, PhD; Sofoklis Kyriazakos[6], PhD; Sadia Anwar[6], MSc; Massimo Corbo[1], MD

[1]Department of Neurorehabilitation Sciences, Casa Cura Policlinico, Milano, Italy

[2]Chair of Wirtschaftsinformatik, esp. Systems Development, Faculty of Business and Economics, Technische Universität Dresden, Dresden, Germany

[3]Neurology Department, Neurodegenerative Diseases Group, Biocruces Bizkaia Health Research Institute, Hospital Universitario Cruces, Barakaldo, Bizkaia, Spain

[4]Ikerbasque: The Basque Foundation for Science, Bilbao, Spain

[5]Universitatea de Medicina si Farmacie “Carol Davila”, Bucuresti, Romania

[6]Department of Business Development and Technology, Aarhus University, Aarhus, Denmark

Corresponding Author:
Peppino Tropea, PhD
Department of Neurorehabilitation Sciences
Casa Cura Policlinico
via Giuseppe Dezza, 48
Milano, 20144
Italy
Phone: 39 024859 ext 3124
Fax: 39 0248593312
Email: p.tropea@ccppdezza.it

Abstract

Background: In the last few years, several studies have focused on describing and understanding how virtual coaches (ie, coaching program or smart device aiming to provide coaching support through a variety of application contexts) could be key drivers for health promotion in home care settings. As there has been enormous technological progress in the field of artificial intelligence and data processing in the past decade, the use of virtual coaches gains an augmented attention in the considerations of medical innovations.

Objective: This scoping review aimed at providing an overview of the applications of a virtual coach in the clinical field. In particular, the review focused on the papers that provide tangible information for coaching activities with an active implication for engaging and guiding patients who have an ongoing plan of care.

Methods: We aimed to investigate the use of the term *virtual coach* in the clinical field performing a methodical review of the relevant literature indexed on PubMed, Scopus, and Embase databases to find *virtual coach* papers focused on specific activities dealing with clinical or medical contexts, excluding those aimed at surgical settings or electronic learning purposes.

Results: After a careful revision of the inclusion and exclusion criteria, 46 records were selected for the full-text review. Most of the identified articles directly or indirectly addressed the topic of physical activity. Some papers were focused on the use of virtual coaching (VC) to manage overweight or nutritional issues. Other papers dealt with technological interfaces to facilitate interactions with patients suffering from different chronic clinical conditions such as heart failure, chronic obstructive pulmonary disease, depression, and chronic pain.

Conclusions: Although physical activity is a healthy practice that is most encouraged by a virtual coach system, in the current scenario, rehabilitation is the great absentee. This paper gives an overview of the tangible applications of this tool in the medical field and may inspire new ideas for future research on VC.

(J Med Internet Res 2019;21(10):e12805) doi:10.2196/12805

KEYWORDS

virtual coaching; rehabilitation; clinical medicine; review; embodied conversational agent; physical activity; health behavior

Introduction

Background

Virtual coaches are seen as key drivers for health promotion in home care settings. Concurrently, against the background of demographic development, the continuity of care from the professional environment into the home has become critical to ensure patients' quality of life and to optimize the economics of medical and social treatments. Nowadays, 1 out of 6 people in the European Union (EU) has a disability [1]. Over one-third of people older than 75 years old have disabilities that restrict them to some extent [2]. Acute disease episodes such as stroke, neurodegenerative diseases, heart attack, or diabetes cause most long-term disabilities and entail very costly care processes. In industrialized countries, stroke is the third major cause of death and the most frequent reason for neurological disability [3]. Rehabilitation is a process by which a patient follows a care plan, first in a protected environment (ie, clinic or rehabilitation center) and second at home, the latter case requiring a tutor. The current evidence shows that home-based programs are effective [4], especially in terms of exercise capacity and health-related quality of life, offering comparable benefits to hospital-based programs [5].

As there has been enormous progress in the fields of artificial intelligence (AI) and data processing in the past decade, the use of virtual coaches has gained increasing attention in the considerations of medical innovations. The term virtual coach refers to a coaching program or smart device aimed at providing coaching support through a variety of applications. The virtual coach will be able to personalize and adapt goals according to the progress achieved by the user in their recovery from their impairments or disabilities. It is broadly defined as any form of coaching using electronic media, including or excluding input from a real (human) coach. The virtual coaches guide and train users through a set of tasks, with the aim of supporting positive actions or assisting in learning new skills. For example, virtual coaches could help users to define and preserve a fitness program, suggest problem-solving skills training, or advise patients with specific medical conditions [6-8].

The latest technological developments have shown promise as effective, accessible, and cost-effective solutions. Recent research about the use of embodied conversational agents (ECAs; ie, a computer-generated cartoonlike character, with the ability to produce and respond to verbal and nonverbal communication) has shown that users can efficaciously form a working alliance relationship with a nonhuman agent (ie, a virtual coach) [9].

Virtual coaching (VC) is becoming increasingly important in medical care and in health-related research investments. For example, the recent EU's Horizon 2020 research framework program funded several projects on VC through a specific call entitled "Personalised coaching for well-being and care of people as they age" [10]. The primary objective of this call was the development of VC solutions that provide personalized advice, guidance, and follow-up support for key age-related issues of daily life that sustain a person's abilities to remain active and independent (eg, diet, physical activity, risk avoidance, and overall wellness).

Objectives

Currently, VC for rehabilitation support in the home environment is still in its initial phases of development from the technical, organizational, medical, and ethical points of view. It is of the foremost importance to provide a terminological and conceptual foundation for this new field. Therefore, we sought to investigate the use of the term virtual coach in the clinical field by performing a methodical review of the relevant literature on VC. Pretests of the body of literature for finding a suitable review method revealed a variety of VC approaches, conversational agents and avatar applications, study designs, diseases, and outcome measures, such that the approach of a traditional systematic review appeared inappropriate. We decided to use the scoping review method [11] to map the key concepts behind a research area as well as the main sources and types of evidence that are available [12]. In addition, we wished to underline that, during the investigative process, we focused on coaching activities intended as innovative modalities designed to engage and guide patients who have an ongoing plan of care.

This review can be used as a starting point by clinical researchers and clinicians interested in VC, revealing the potential of VC systems in the clinical domain, and supporting the understanding of the essential components of such interventions.

Methods

Search Strategy

We chose to adopt and follow the Arksey and O'Malley scoping review methodology [11]. The literature review was conducted on the PubMed [13], Scopus [14], and Embase [15] databases to find VC papers focused on the medical domain. Studies were collected up to July 2019.

Considering the main key concept for our research, specifically a coaching program or smart device aimed at providing coaching support through a variety of applications, we identified the relevant key words (namely, virtual coach and virtual trainer) for an initial search strategy. These keywords were then refined and used for the identification of related terms. Thus, the final keywords were as follows (syntax was adjusted for each database as necessary): "virtual coach" (OR "virtual coach interventions" OR "virtual coaching"); "virtual trainer"; "virtual training" (OR "virtual training platform" OR "virtual training system"); "virtual therapist"; "virtual nurse"; "virtual agent"; "embodied conversational agent"; "avatar assisted therapy" (OR "avatar mediated training" OR "avatar therapy"); "e coach."

In addition, wildcard symbols, such as hyphens or inverted commas, were used to consider all possible variations of root words (eg, plural). Limiting categorical area terms were used for the Scopus database: *Medicine*, *Health Professions*, and *Psychology*. Only published journal papers written in English were included, but there was no limitation regarding the publication date.

In addition to the electronic database search, a manual search was performed in JMIR for issues dated up to July 2019. Moreover, we carefully checked references to each article we found to avoid missing relevant papers not identified in the electronic search (namely, handsearching). Two authors (PT and IS) conducted the literature search, and they stored the references in a spreadsheet program, removing duplicates.

Inclusion and Exclusion Criteria

In a second step, after electronic searching and handsearching, we used the titles and abstracts of the identified articles to check their pertinence to the medical domain. At this step, we selected papers that focused on specific VC activities dealing with clinical or medical contexts, excluding those aimed at surgical settings or electronic learning (e-learning) purposes.

The preselected abstracts were considered for full-text analysis if the related VC had a reasonable sense of agency, with some extent of autonomous behavior able to exhibit some form of reasoning. Furthermore, 3 reviewers (PT, IS, and MCo) independently examined the search results of possible relevant papers. Consensus rounds were used to solve disagreements in the selection.

Full-Text Assessment

Publications that satisfied the initial inclusion and exclusion criteria were downloaded into bibliography manager software [16] for further screening toward a deeper investigation of the same, previous listed inclusion and exclusion criteria. To meet the objectives of this scoping review, we gathered the following information from each paper: (1) clinical field of application, (2) interface used, (3) study typology, (4) type and number of enrolled participants, (5) intervention duration, (6) study aim or purpose, (7) outcome measurement, and (8) overall conclusions.

Data were summarized in a descriptive form. A quantitative assessment was attempted, but the wide heterogeneity in the study design, instrumental technologies, characteristics of the studied subjects, and analytical methods adopted in the studies prevented any type of meta-analysis of data.

Results

Quantitative Results and Overview

Figure 1 shows the paper selection process. A total of 294 records were identified in PubMed [13], 264 in Scopus [14], and 271 in Embase [15] databases. Overall, 398 nonduplicate records were identified in the initial electronic search (274 were duplicated, 154 were common to all database, 47 were unique from PubMed, 56 were unique from Scopus, and 19 were unique from Embase).

After reading the titles and the abstracts, 309 records clearly did not meet our eligibility criteria; the 89 remaining records were selected to undergo further analysis. Several articles were related to clinical aspects; however, they were not focused on specific coaching activities. Therefore, we further excluded 49 records that did not meet the inclusion and exclusion criteria for full-text assessment. In detail, 27 papers were excluded because they were related to mere virtual training approaches, without any implication on coaching activities or clinical applications, 8 analyzed e-learning platforms, and 6 were related to emotional aspects of ECA (nondirectly related to clinical interaction). In addition, 3 were excluded because they presented a model of user-interface communication services, 2 were only related to reminder or agenda services, 2 presented the analysis of interactions between users and commercial virtual environments, and 1 was a review on behavioral changes. In addition, using handsearching, we retrieved 6 additional records, 1 of which was a conference paper previously not included by automatic research.

After a careful revision of inclusion and exclusion criteria, a total of 49 records were finally selected to undergo the full-text review. In addition, 16 of the 46 studies selected were clustered as randomized controlled trials (RCTs), 21 as exploratory studies, 5 as review articles, and 5 as theoretical studies. Figure 2, showing the distribution in time of the selected 46 scientific articles in terms of their publication date (ranging from 2005 to July 2019), reflects a noteworthy increase in interest regarding the use of VC in the clinical setting in the last few years.

Figure 1. Flowchart describing study's identification and selection.

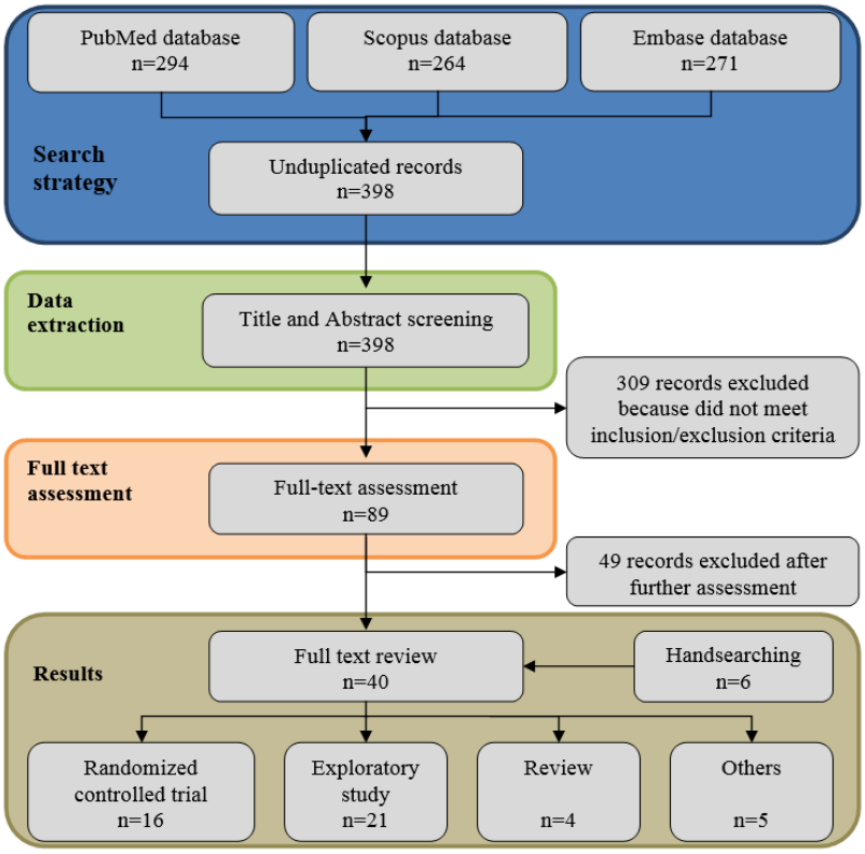
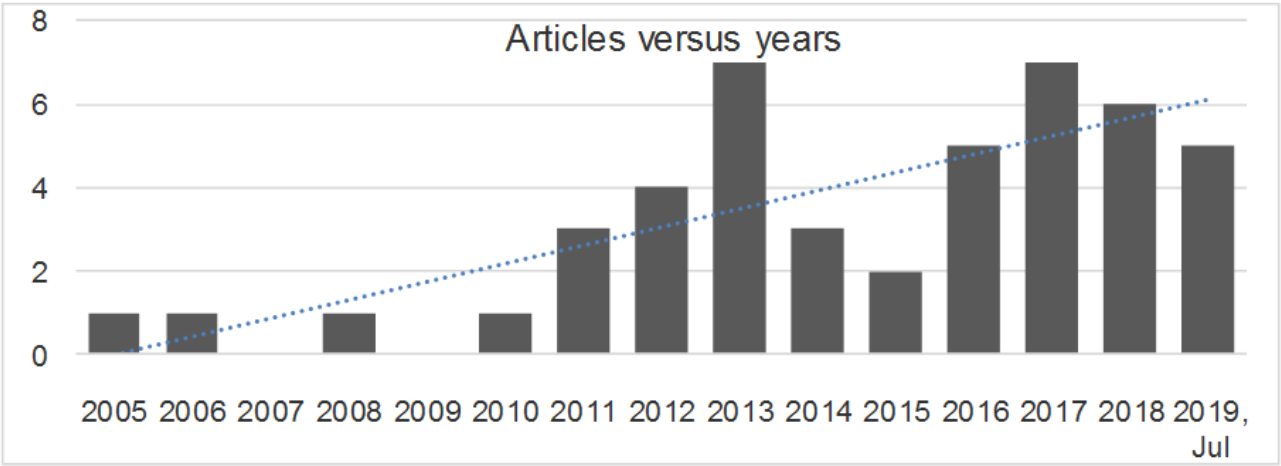


Figure 2. Included articles versus years.



Tables 1-5 summarize the main features of the identified scientific articles on VC.

Table 1. Studies on virtual coaching applied to clinical or medical contexts included in the analysis as randomized controlled trials.

Paper reference	Clinical field of application	Interface	Subjects enrolled	Subjects (n)	Testing period (in weeks, unless otherwise specified)
Allen et al 2008 [17]	Chronic conditions	Not real-time text message (email)	Chronic pain, depression, or impaired mobility	121	4
Gabriele et al 2011 [18]	Obesity in adults	Not real-time text message (email)	Healthy overweight	104	12
Martorella et al 2012 [19]	Chronic conditions	ECA[a] (virtual nurse)	Adults scheduled for their first cardiac surgery	60	1 day
Watson et al 2012 [20]	Obesity in adults	ECA with verbal and nonverbal communication	Healthy overweight or obese	70	12
Bickmore et al 2013 [21]	Nutrition and physical activity	ECA (computer-based)	Healthy	122	8
Bickmore et al 2013 [22]	Physical activity	ECA with verbal and nonverbal communication	Healthy elderly	263	8 (follow-up 1 year)
Wijsman et al 2013 [23]	Physical activity	Avatar	Healthy elderly	235	12
Friederichs et al 2014 [24]	Physical activity	Avatar (textual interface)	Healthy	958	8
Vroege et al 2014 [25]	Physical activity	Avatar	Healthy elderly	235	12
Broekhuizen et al 2016 [26]	Physical activity	Avatar	Healthy elderly	235	12
Leahey et al 2016 [27]	Obesity in adults	Periodic email contact	Healthy overweight	75	40
Ritchie et al 2016 [28]	Chronic conditions	Interactive voice and a dashboard for the nurse to review data	Chronic heart failure and COPD	511	12
Gardiner et al 2017 [29]	Health promotion	ECA	Healthy	61	4
King et al 2017 [30]	Physical activity	ECA (computer-based)	Healthy elderly	245	52
Cotè et al 2018 [31]	Chronic conditions	ECA (virtual nurse)	Kidney transplant patients on immunosuppressive medication	70	1 day
Hui et al 2018 [32]	Physical activity	ECA with verbal; face-to-face communication	Healthy	200	12

[a]ECA: embodied conversational agent.

Table 2. Studies on virtual coaching applied to clinical or medical contexts included in the analysis as randomized controlled trials.

Paper reference	Study aim	Outcome measurement	Conclusion
Allen et al 2008 [17]	Engage and empower patients to collaborate with their primary care physician in managing their health conditions	Access to Web site; access to personal online worksheets; number of emails sent to the electronic coach	Nurses can play an important role, joining efforts to develop new territory to promote patients as partners in managing their health conditions.
Gabriele et al 2011 [18]	The effects of electronic coach support on weight loss, dietary behavior, physical activity, and engagement	Waist and hip circumstances; questionnaire for intervention engagement; Social Support Inventory; dietary behavior	Nutrition education apps are feasible and acceptable solutions to support health promotion interventions.
Martorella et al 2012 [19]	Investigate the preliminary effects of a virtual nursing intervention to improve pain relief in patients undergoing cardiac surgery	Questionnaires; pain intensity score (at the time of admission and after surgery)	Web-tailored approach can increase accessibility to health education and promote pain relief without generating more costs.
Watson et al 2012 [20]	Increase activity levels, via step count, in overweight or obese individuals	Step count	The virtual coach was beneficial in maintaining activity level.
Bickmore et al 2013 [21]	Promote both physical activity and fruit and vegetable consumption through a series of simulated conversations with users on their home computers	IPAQ[a]; step count; National Institutes of Health - National Cancer Institute (NIH-NCI) fruit and vegetable scan; weight	Automated health intervention software designed for efficient re-use is effective at changing health behavior.
Bickmore et al 2013 [22]	Compare the efficacy of a computer-based physical activity program with that of a pedometer control condition in sedentary older adults	Daily step count	An automated exercise promotion system deployed from outpatient clinics increased walking among elderly over the short-term.
Wijsman et al 2013 [23]	Assess whether a Web-based intervention increases physical activity and improves metabolic health in inactive older adults.	Wrist activity monitor; anthropometric measures; blood samples analysis	In inactive older adults, a 3-month Web-based physical activity intervention was effective in increasing objectively measured daily physical activity and improving metabolic health.
Friederichs et al 2014 [24]	Determine whether a Web-based physical activity intervention based on motivational interviewing with an avatar results in more positive appreciation and higher effectiveness of the intervention, when compared to an intervention that is purely text-based	Web-based questionnaires; IPAQ	Avatars that do not strengthen the social relationship with the user do not enhance the intervention impact.
Vroege et al 2014 [25]	Assess how many participants successfully reached the physical activity level as targeted and the effects of the intervention on body composition and metabolic health	Wrist activity monitor; body mass index; blood samples analysis	Of the intervention group, 42.0% reached their daily physical activity end goal, which was associated with a markedly better effect on body composition and metabolic health compared to the effect in the entire intervention group.
Broekhuizen et al 2016 [26]	Assess if an internet-based intervention aimed to increase physical activity was effective in improving quality of life of inactive older adults	Wrist activity monitor; Dutch paper version of the Research and Development 36-item health survey	Internet-based physical activity program was effective in improving quality of life in 60-70-year-olds after 3 months.
Leahey et al 2016 [27]	Efficacy of a novel approach to weight loss maintenance based on modifying the cost–benefit ratio	Weight	Internet delivered cost–benefit approach to weight loss maintenance may be effective for long-term weight control.
Ritchie et al 2016 [28]	Evaluate the impact of a technology-supported care transition support program on hospitalizations, days out of the community and mortality	Number of rehospitalizations, death, days in the hospital, and out of the community	Clinically meaningful reduction in 30-day rehospitalization rates in chronic obstructive pulmonary disease patients when using an interactive voice response–enhanced care transition intervention.
Gardiner et al 2017 [29]	Evaluate the feasibility of using an ECA to teach lifestyle modifications to urban women	Patient Health Questionnaire; Household Food Insecurity Access Scale; Self-Efficacy for Exercise Scale; Perceived Stress Scale	It is feasible to use an ECA to promote health behaviors on stress management and healthy eating among diverse urban women.
King et al 2017 [30]	Compare the effectiveness of 2 linguistically and culturally adapted, community-based physical activity interventions with the potential for broad reach and translation	Changes in walking and other forms of physical activity measured via self-reporting and accelerometers.	The intervention has substantial potential to reduce the health disparities gap by influencing a key health behavior in underserved populations.

Paper reference	Study aim	Outcome measurement	Conclusion
Cotè et al 2018 [31]	Evaluate the acceptability, feasibility, and preliminary efficacy of virtual nurse intended to support medication adherence among kidney transplant recipients	Web-Based Nursing Intervention Acceptability Scale; number of completed sessions; medication adherence measures	The results support the feasibility and acceptability of proposed virtual nurse. It could constitute an accessible adjunct in support of existing specialized services.
Hui et al 2018 [32]	Evaluate the effectiveness of an information technology-based lifestyle intervention program on improving physical activity level and health status in a sample of middle-aged Hong Kong adults	Physical Activity measured by accelerometer; IPAQ	The information technology–based lifestyle intervention is fast, inexpensive, flexible, and convenient for adults, especially those with a busy work life.

[a]IPAQ: International Physical Activity Questionnaire.

We identified 5 reviews on VC. To define the fundamentals and elementary components of the term virtual coach, we first analyzed the reference paper by Ding et al [53]. Recently, another review examined electronic health (eHealth) interventions with the combined use of self-tracking and persuasive electronic coaching (eCoaching) [54]. In the psychiatric field, 2 reviews were focused on the use of ECA in psychotherapy [55] and the use of conversational agents for benefits in psychoeducation and self-adherence in the screening, diagnosis, and treatment of mental illnesses [56]. Another recent review identified some technological systems that could rehabilitate elderly patients with insomnia, including virtual coaches [57].

Regarding their clinical fields of application for VC, most of the identified articles classified as RCTs directly or indirectly addressed the topic of physical activity [22-26,30,32]. Some RCTs were mainly focused on the use of VC to manage being overweight [18,20,27] or nutritional issues [21] instead of physical activity. Other RCTs dealt with technological interfaces to facilitate interactions with patients suffering from different chronic clinical conditions such as heart failure, chronic obstructive pulmonary disease (COPD) [28], depression and chronic pain [17] or therapy adherence [19,31]. Similar usage contexts, namely physical activity [33,34,37,40,42,48], nutrition [6], and chronic condition [35,39,41,43,47,51,52], were also found in most exploratory studies. Finally, we also identified other interesting exploratory studies that evaluated the psychological effects [36,44] and the psychoeducational benefits of VC for the users [38,45,46].

Conversely, we identified several scientific articles that highlighted the growing interest in the use of computer-based ECAs for clinical purposes, both in the context of physical activity [19-22,30-32,35,40,41,43,51] and psychological applications [36,38,42,44,45,49] (Figure 3). Moreover, the ECAs that simulated the key features of a human conversation have also been the subject of a scoping review centered on clinical psychology [55] and 5 theoretical articles focusing on their modeling and architecture [58-62]. Figure 4 shows the distribution of the clinical conditions treated in the 46 studies.

Table 3. Studies on virtual coaching applied to clinical or medical contexts included in the analysis as exploratory studies.

Paper reference	Clinical field of application	Interface	Subjects enrolled	Subjects (n)	Testing period (in weeks, unless otherwise specified)
Guillen et al 2005 [33]	Physical activity	Touch screen monitor	Healthy	30	8
Segerståhl et al 2011 [34]	Physical activity	Not real-time text message (Web-based)	Healthy	30	3
Cotè et al 2012 [35]	Chronic conditions	ECA[a] (virtual nurse)	HIV positive	71	1 day
Novielli et al 2012 [36]	User's reactions	ECA with verbal; face-to-face communication	Healthy young	30	Variable
Ellis et al 2013 [37]	Physical activity	ECA (computer-based)	Parkinson disease	20	4
Hudlicka 2013 [38]	Meditation Training	ECA	Healthy	32	3
Kreps et al 2013 [39]	Chronic conditions	SMS text reminders	Crohn disease	30	35
Silveira et al 2013 [40]	Physical activity	ECA with verbal and nonverbal communication (tablet based)	Healthy elderly	44	12
van Vuuren et al 2014 [41]	Chronic conditions	ECA with verbal and nonverbal communication (virtual therapist)	Chronic aphasia resulting from a single left hemisphere stroke	8	9
Adams et al 2015 [42]	Physical activity	Not real-time text message (email) and phone call	Coronary artery bypass grafting	1	16
Cotè et al 2015 [43]	Chronic conditions	ECA (virtual nurse)	HIV positive	26	__[b]
Stevens et al 2016 [44]	User and ECA interactions	ECA (computer-Based)	Healthy young	40	1 day
Tielman et al 2017 [45]	Psycho-education	ECA	Healthy	46	3 days
Chi et al 2017[46]	Social isolation in older adults	Pet avatar (tablet based)	Healthy elderly	10	12
Gabrielli et al 2017 [6]	Nutrition	Mobile app	Overweight children and their parents	6 children 6 parents	6
Klaassen et al 2018 [47]	Chronic conditions	Text messages and notifications to the website and to the application	Adolescents with type 1 diabetes	21	6–8
Oyibo et al 2018 [48]	Physical activity	Video	Healthy	659	1 day
Richards et al 2018 [49]	Urinary incontinence	ECA (Web-based)	Children	74	26
Suganuma et al 2018 [50]	Psychotherapeutic intervention	ECA with verbal; face-to-face communication	Healthy	128	2
Dworkin et al 2019 [51]	Chronic conditions	ECA	HIV positive	43	12
Srivastava et al 2019 [52]	Chronic conditions	Virtual coaching	Patients with prediabetes	10	1 day

[a]ECA: embodied conversational agent.

[b]Missing data.

Table 4. Studies on virtual coaching applied to clinical or medical contexts included in the analysis as exploratory studies.

Paper reference	Study aim	Outcome measurement	Conclusion
Guillen et al 2005 [33]	Assess technological platform and the fitness condition.	Heart rate; VO_2	The coach is based on the personalized training programs adapted to user's characteristics and preferences and on a continuous assessment of the actual fitness status.
Segerståhl et al 2011 [34]	Investigate how users incorporate a system employing 2 modes of delivery into their training and analyze benefits in personal exercise monitoring.	Heart rate	Personal exercise monitoring systems may be improved by more systematically combining mobile and Web-based functionality.
Cotè et al 2012 [35]	Study of the acceptability and feasibility of a Web application which was designed to empower people living with HIV to manage their daily antiretroviral therapies	Acceptability questionnaires; field notes; observations	The results of the study support the feasibility and acceptability of the intervention.
Novielli et al 2012 [36]	Investigate the user's reactions to received suggestion by an embodied conversational agent (ECA) playing the role of artificial therapist in the healthy eating domain	Classification of speech; user's reactions	The adaptation of the dialogue is crucial for an effective persuasive interaction.
Ellis et al 2013 [37]	Feasibility, acceptability, and preliminary evidence of effectiveness of a virtual exercise coach to promote daily walking in community dwelling persons with Parkinson disease.	Six-minute walk test; gait speed; number of steps per day; retention rate; satisfaction; interaction history	Sedentary persons with PD successfully used a computer and interacted with a virtual exercise coach. Retention, satisfaction, and adherence to daily walking were high over 1 month.
Hudlicka 2013 [38]	Develop and evaluate a virtual mindfulness coach for training and coaching in mindfulness meditation	Web-administered surveys 5-point; Likert scale	Virtual coach-based training of mindfulness is both feasible, and potentially more effective, than a self-administered program.
Kreps et al 2013 [39]	Describe how electronic health communication programs can be improved by using artificial intelligence to increase immediacy	Weight; patient's activity; sleep patterns	Artificial Intelligence can enhance the ''immediacy'' of eHealth by humanizing health promotion efforts, promoting physical and emotional closeness.
Silveira et al 2013 [40]	Investigate which information technology-mediated motivation strategies increase adherence to physical exercise training plans in older people	Adherence and attrition; gait speed; motivation instruments	The social motivation strategies were more effective to stimulate the participants to comply with the training plan and remain on the intervention.
van Vuuren et al 2014 [41]	Analysis of performance of a system for delivering speech and language therapy to people with aphasia, delivered by a virtual therapist	Performance in terms of word accuracy	For persons with aphasia, receiving treatment in an ecologically valid real-world setting delivered by a virtual therapist that provides more cues than not can lead to faster learning.
Adams et al 2015 [42]	Investigate how modify desired movements and activities in a way that minimizes shoulder joint abduction, extension, and flexion	Blood pressure; amount of work done in weight training	The VC, based on a sport-specific, symptom-limited exercise program, would enable the patient to train at a higher intensity than is typically allowed.
Cotè et al 2015 [43]	Explore and describe how patients living with HIV experience receiving customized asynchronous accompaniment via a virtual nurse	Semistructured interviews to get participants to share their experience of the intervention through personal stories and what they thought and felt during their participation	The virtual nurse humanized the experience and helped them acquire new skills for achieving optimal antiretroviral therapy adherence.
Stevens et al 2016 [44]	Investigate the effect of features of human behavior on the quality of interaction with an ECA	Number of errors in ECA speech; multiple-choice questions	Influences from mimicry can be explained by visual and motor simulation, and bidirectional links between similarity and liking.
Tielman et al 2017 [45]	Study the preferable presentation mode for improving adherence.	Behavioral data; questionnaires	Both the attitude towards the virtual agent and how well the psychoeducation was recollected were positively related to adherence in the form of task execution.
Chi et al 2017 [46]	Examine the perceived acceptance and utility of a tablet-based ECA (termed *digital pet*) for older adults	Semistructured, individual interviews for testing of a digital pet companion	A digital pet can provide older adults with companionship and enhance social interaction.

Paper reference	Study aim	Outcome measurement	Conclusion
Gabrielli et al 2017 [6]	Describe the design and development of a nutrition education app and the results of a formative evaluation with families	Knowledge of the Mediterranean diet; URICA (University of Rhoda Island Change Assessment)-short-form scale; intention to use technology for nutrition education	The user-centered design showed that nutrition education apps are feasible and acceptable solutions to support health promotion interventions in primary care.
Klaassen et al 2018 [47]	See how patients with type 1 diabetes experience the game with a VC	System Usability Scale; semistructured interview to explore user experiences	User evaluations with patients under pediatric supervision revealed that the use of mobile technology in combination with Web-based elements is feasible.
Oyibo et al 2018 [48]	Evaluate determinants of bodyweight exercise performance in the context of behavior modeling in fitness apps	Social cognitive theory instruments	The study provides a set of guidelines for the design of persuasive technologies for promoting regular exercise behavior.
Richards et al 2018 [49]	Evaluate a novel approach that involves a website and virtual specialist for patients while they are awaiting their specialist appointment	Cross-cultural continence-specific quality of life; Pediatric Incontinence Questionnaire; ehealth literacy survey (eHEALS) ; Newest Vital Signs test; Working Alliance Inventory	A novel approach that involves a website and virtual specialist for patients while they are awaiting their specialist appointment showed an overall improvement in 74% of patients with urinary incontinence.
Suganuma et al 2018 [50]	Use an internet-based cognitive behavioral therapy preventative mental health measure	World Health Organization-Five Well-Being Index; Kessler 10; Behavioral Activation for Depression Scale	The internet-based cognitive behavioral therapy with the embodied conversational agent can be used in mental health care.
Dworkin et al 2019 [51]	Evaluate the feasibility, acceptability, and preliminary efficacy of an ECA intervention to improve adherence to antiretroviral therapy	Adherence; acceptability; feasibility, pre- versus posthealth literacy; pre- versus post-self-efficacy	The pilot study of demonstrated acceptability and preliminary efficacy in improving adherence in this important population.
Srivastava et al 2019 [52]	Evaluate performance relative to behavior stages associated with long-term behavior modification	Continuity in adherence to the program	The strength of the physician–patient relationship appears to allow people with prediabetes to skip or advance rapidly through behavioral stages in the process of lifestyle modification.

Table 5. Reviews and theoretical studies on virtual coaching applied to clinical or medical contexts.

Paper reference	Clinical field of application	Interface	Study typology
Ding et al 2010 [53]	__[a]	Overview of virtual coach interventions	Review
Lentferink et al 2017 [54]	Lifestyle	e-Coach[b]	Review
Provoost et al 2017 [55]	Psychotherapeutic interventions in clinical psychology	ECA[c] (virtual human characters on computer screens to robots) and communication	Review
Vaidyam et al 2019 [56]	Psychiatry	Conversational agent, chatbot	Review
Salvemini et al 2019 [57]	Insomnia and sleep disorders	Virtual coach	Review
de Rosis et al 2006 [58]	Health promotion	ECA with verbal; face-to-face communication	Theoretical study
Cotè et al 2011 [59]	HIV positive	ECA (virtual nurse)	Theoretical study
Perez et al 2016 [60]	User and ECA interactions	ECA	Theoretical study
Cotè et al 2017 [61]	Chronic conditions	ECA (virtual nurse)	Theoretical study
Fadhil et al 2019 [62]	Health promotion	ECA	Theoretical study

[a]Not applicable.

[b]e-coach: electronic coach.

[c]ECA: embodied conversational agent.

Figure 3. On left: number of studies (randomized controlled trial or exploratory) versus interface modalities. On right: some examples of embodied conversational agents (ECAs) reported in the selected articles: (a) Friederichs et al 2014, (b) Tielman et al 2017, (c) Stevens et al 2016, (d) de Rosis et al 2006, Novielli et al 2012, (e) Bickmore et al 2013a, Ellis et al 2013.

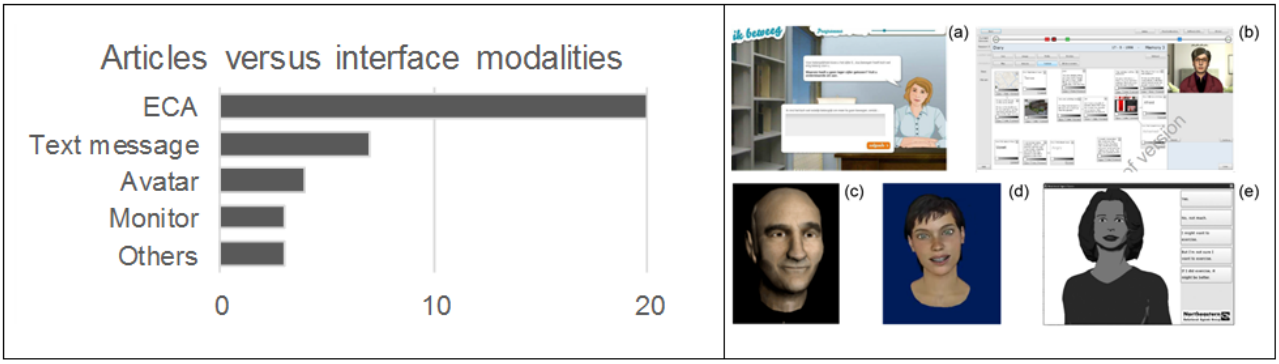
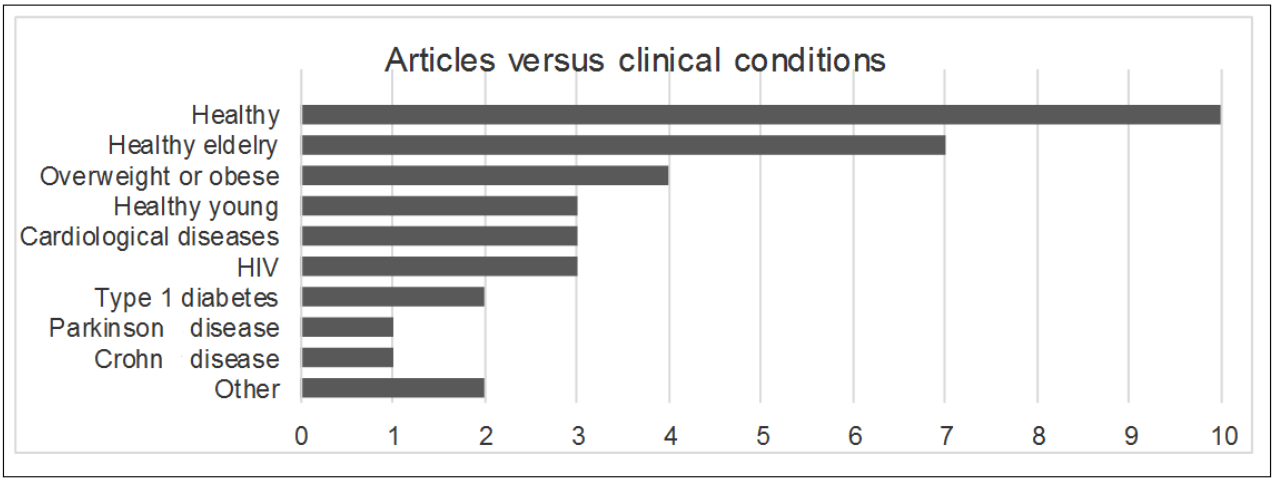


Figure 4. Number of studies (randomized controlled trial or exploratory) versus clinical conditions.



Definition of Virtual Coach

Overview

The term virtual coach was firstly used in the athletic context during the 1950s, in reference to sporting coaches that were leading the teams behind the scenes or over distance. The earliest medically oriented usage of Virtual Coach was in 1997 in an operating room for guiding a nonstandard surgical procedure. Nowadays, the term virtual coach is a specific technological concept referring to *ad hoc* platforms comprised of one or more devices with a dedicated software that are used for training and coaching interventions. VC platforms use several tasks to support the learning of new activities and encourage positive or correct behavior. According to Ding et al, 4 components (ie, self-monitoring, context awareness, coaching strategy, and interface modality) define the space of action of such interventions and address when, how, and what message, feedback, and stimuli to deliver.

Self-Monitoring

This could be defined as the observation and recording of a specific activity or a specific condition, related to personal and physical information (eg, activity levels, calorie intake, and weight). Self-monitoring is performed (also in real time) with a wide variety of sensors that allow the tracking of the user performance, triggering messages to the user only when they are necessary and relevant.

Context Awareness

This consists of a collection of specific contextual parameters (eg, location, time, identity, and activity) and of the responsiveness to an interruption of a task based on that contextual information. However, many more factors, such as the emotional state of the user and the reason and modality of the interruption, are crucially linked to the responsiveness [63]. Sensor technologies combined with VC algorithms empower a system to improve the identification of contextual elements (eg, location and user activity), optimizing the identification of activity discontinuation by the user and delivering prompt messages in appropriate moments.

Coaching Strategy

A coaching strategy selects which kind of content should be contained in each message delivered to the user, distinguishing general versus specific messages and the incorporation of proper emotional effects within them. The strategy also aims to perform a dynamic adaptation of the activities to the needs, performance, and improvements of the user, which are collected through the two self-monitoring and context awareness phases. Due to the adaptation, the annoying effect could be avoided and the user may be more confident in their use of the system. Moreover, positive messages may help to reinforce the performance, comfort, and interest in using the system. Conversely, the use of negative feedback and messages should be limited as they are normally inappropriate and the aim is to decrease errors.

Interface Modality

This includes the different kinds of interfaces (ie, visual, auditory, and tactile) that can be used to deliver coaching messages to users. As matter of fact, ECAs have become the most popular interface because of increasing innovations in the field of computer graphics. ECA interfaces could be very different, ranging from simple animated characters to humanlike or animal-like entities, talking animals, or many other varying forms. One of the common trends is to design animated humanlike agents and then characterize them with a great heterogeneity of features, from its gender, aesthetics, and voice to its personality and behavioral patterns.

Interaction Between Virtual Coach and the User

As previously mentioned, the ability to provide coaching messages in a way that is personalized both to the user and the situation will likely develop a trustworthy relationship between the VC platform and the user, improving the acceptance and effectiveness of the technology. In this context, the interaction between the virtual coach and the user acquires a priority role. From the analyses of the literature, a description of the interaction between VC and users was found as the relevant topic. Several articles examined this theme in depth [18,20,21,36,49,53,55]. The most comprehensive description of the interaction between VC and users was proposed by Watson et al [20]. In their study, they suggested several steps to organize VC-user communication as a structured process: (1) opening with salutation and social contact; (2) proceeding to review of data acquired, feedback, and goal setting; (3) advice about specific topics (eg, activities and diet); (4) guarantee to time of next contact; and (5) ending with encouragement and farewell.

Moreover, the structure and format of the dialogue, together with the content of individual expressions, needs to be personalized according to the users' progress in the system, their current status, and the speech context [20]. An acceptable interface for communication should be able to distinguish the communicative and social attitudes of users, and adjust the dialogue and the interaction style accordingly [36]. Due to this personalization, individuals who did not reach their proposed objectives or who had a different attitude towards the communication would have had a different interaction than those who met their goals. The ability to provide personalized feedback and support, in an interactive manner, was the most successful coaching feature [38].

According to Kreps et al [39], the success of health communication interaction depends upon multiple factors, including the accuracy, timeliness, fidelity, persuasiveness, and sensitivity of messages exchanged. In addition, the immediacy is a critical factor in determining whether communication processes work and could help users to work together to achieve important health goals [39].

A group of 20 out of 46 revised papers (see Tables 1 and 3 and Figure 3) used the human ECAs as interface modality of VC. Provoost et al [55] published a scoping review about the use of the ECA in the field of clinical psychology. The authors defined the concept of an ECA via the classification of the roles, characteristics, and skills of these interfaces. Thus, the ECA consists of 3 components:

1. the interface that allows users to communicate with the ECA (eg, from Web-based questionnaires to real-time audio and video input);
2. the algorithm that consists of the mental capacities of the ECA, allowing the ECA to react empathically with users;
3. the so-called embodiment (or visual representation) of the ECA, meaning how it appears (eg, humanoid, toylike, photograph, or fantasy). The look of the ECA is also defined by the way in which it communicates with users (eg, verbally vs nonverbally, text messages vs speech, and with or without gestures and facial expressions).

The authors also categorized the way in which information was conveyed: ECA and users could communicate with each other through speech, facial expressions and eye gaze, hand and body gestures, or simply text [55]. The negative experience performed by Friederich et al [24] corroborated the relevance of considering the multiple verbal and nonverbal communication modalities when designing ECAs. In the cited paper, the avatar, with limited relational skills and unable to respond in gestures to the user's state and input, was not able to enhance the intervention and therefore did not increase the intervention's impact.

Simulating facial expressions and eye gaze (face-to-face conversation) is particularly important in health care when dealing with users with low literacy because in those cases, face-to-face encounters with health providers remain the most effective interaction modalities [21]. Regarding the aesthetical characteristics of the ECA, it is important to state that the level of trustworthiness is increased by the personification of the agent. The most recent systems incorporate bots, games, and simulated environments to talking/thinking/teaching ECAs, allowing them to understand the effect of different features of the system on human acceptability, understanding, and learning [44]. The effects of mimicry on likability in human-ECA interaction were tested to investigate whether visual cues displayed appropriate signals than those in the no-mimicking condition. Similarly, the effect of the ECA's facial gestures (ie, smiling, winking, and rolling eyes) was investigated in a parallel experiment to understand whether comprehension of a story was increased when recited by an ECA with facial expressiveness [44]. Expressive facial gestures may aid learning through increased redundancy and increased cognitive load. In contrast with this, it is reported that some users experienced a phenomenon referred to as the uncanny valley [64], where the general trustworthiness of the agent negatively correlates with the degree of visual realism of the ECA form. In other words, ECAs that are more cartoonish are perceived as more acceptable and effectively realistic than more visually realistic ones [38].

Regarding the social aspects of ECAs, it is worth mentioning that one of the selected articles described 4 social roles for ECAs [55]: (1) the social interaction partner aims to engage in an interaction with the user to improve specific social skills, (2) the tutor aims to teach something, (3) the coach aims to motivate and engage the user, and (4) the health care provider aims to simulate the behavior of a health care provider.

In addition, a pilot study evaluated how older adults could interact with a pet avatar [46]. This study examined the perceived acceptance and utility of a tablet-based conversational agent in the form of a pet avatar that was used by older adults over 3 months ,during daily interactions. The results disclosed that this interface (ie, a digital pet) could provide older adults with companionship and enhance social interaction.

In conclusion, studies indicate that ECAs have the following benefits:

1. Social interaction: relations between a human being and a machine are social; an avatar supports contact in a natural way, like with other human beings.
2. User attention: animated characters can attract the user's attention.
3. Real simulation: embodied relational agents provide the illusion of liveliness and interacting with a real person.
4. Body language information: facial expressions of the avatar contain a lot of information.
5. Trustworthiness and believability: the level of trustworthiness is amplified by the humanization of the ECA; a realistic face seems to be rated as more intelligent, engaging, and likeable.

Medical Care Sectors

Summary

The analysis of the literature showed that the application of the VC has involved different areas of the medical domain, which allowed us to address the problem from different angles. In particular, the studies evaluated specific topics, such as physical activity and nutritional regimen, or wider contexts linked to the reduction of the risk of illness. The following are the main topics covered.

Lifestyle: Physical Activity

The physical activity of adults and elderly people is increasingly recognized as an essential lifestyle behavior in maintaining health and preventing diseases. Therefore, it is not surprising that many articles have tried to introduce an automated exercise promotion system to support motor activity in healthy or overweight elderly subjects.

At the beginning of this decade, Segerståhl and Oinas-Kukkonen [34] supported the idea that physical exercise monitoring systems could be improved by combining mobile and Web functionality. So, they conducted a qualitative field study with 30 participants using a heart rate monitor and a Web service in their training for 3 weeks. Despite the limitations of the study because of the lack of a control group, the authors showed that a system employing multiple modes of delivery may influence personal exercise monitoring as perceived by users, and highlighted the importance of future developments of intelligent coaching that can provide useful and personalized information. Indeed, keeping an optimal physical status is a critical issue in an ageing society. Hui et al [32] tested the effectiveness of a Web-based virtual training system with ECAs to improve physical activity in a sample of middle-aged adults. The system was flexible and convenient for adults, especially for those with a busy work life. In a preliminary validation study, Fadhil et al [62] discussed the outcome of a 1-month use of a conversational agent–assisted health coaching system designed to support health intervention delivery to individuals or groups, showing initial promising results in the engagement and adherence of users and the role of a conversational agent in delivering health-promoting activities.

Among the elderly population, Bickmore et al [21,22] evaluated the efficacy of an automated intervention with the ECA to improve physical activity and fruit and vegetable consumption in sedentary older adults, showing an effective result at changing health behavior with this computer-based program. Regarding physical activity, results of the RCT demonstrated that participants who received the ECA intervention walked significantly more steps, recorded by a pedometer, than control participants at 2 months [22]. In particular, the authors highlighted 2 fundamental factors on the success of physical activity promotion: the importance of face-to-face conversation, using voice, between the ECA and the participants, and their level of health literacy. A subsequent work by Friederich et al [24] confirmed that the relational skills of the avatar are very important in strengthening the relationship with the user and may enhance intervention impact. In fact, they developed a Web-based physical activity intervention based on motivational interviewing that was done by an avatar positioned behind a desk in a small office (Figure 3, on right). The avatar communicated through text balloons, without the use of voice or nonverbal expressions such as empathic gestures and eye and head movements. Consequently, the avatar used in the study was not able to increase intervention effectiveness and the authors concluded that the action of a virtual coach with more complex relational skills had to be recommended for future research. In this context, King et al [30] conceived a study design to evaluate the comparative effects of physical activity recommendations by humans or computers in unserved populations. The authors, confident of the results obtained by using a virtual advisor (ie, ECA) in a preliminary study [65], hoped to obtain, through a comparative effectiveness trial, the confirmation that this program could offer a substantial potential to reduce the health disparities gap by influencing a key health behavior in underserved populations.

In addition, a Web-based physical activity intervention was proposed to increase health and reduce risk factors in the elderly population. In several studies from the same group [23,25,26], an internet program aimed to increase physical activity and improve quality of life of inactive older adults using monitoring and feedback by accelerometers and digital coaching. The authors found that an internet-based physical activity program was effective in improving quality of life after 3 months, particularly in participants that reached their individually targeted increase in daily physical activity. More importantly, they also observed an improvement of objectively measured daily physical activity and metabolic health after a 3-month, Web-based, physical activity intervention. Such Web-based interventions provide new opportunities for large-scale prevention of chronic and metabolic diseases in the aging population.

Similarly, in a preclinical exploratory trial [40], an information technology–based training app that runs on tablets and monitors

older people was used to follow personalized training plans autonomously at home, and to improve balance and strength in 44 older adults. The app seemed to assist and motivate the subjects to autonomously perform strength-balance exercises and increase their gait speed. The social motivation strategies were more effective at stimulating the participants to comply with the training plan and lowered the dropout rate.

Lifestyle: Nutrition and Obesity

Watson et al [20] demonstrated a sustained level of activity in overweight adults provided with a VC in addition to a pedometer and Web-based feedback, compared with a decline seen in those without VC. Even the use of an electronic coach (e-coach) to support and help users has resulted in them keeping their weight under control [18,27], or the development of an intelligent virtual agent (ie, ECA) to persuade people to improve their eating habits [36] has provided promising results. Limited experiences are documented in the context of nutrition education apps, which however seem to be potentially useful in providing effective prevention and health promotion programs for overweight children [18].

To provide guidelines for the design of persuasive technologies for promoting regular exercise behavior, Oyibo et al [48] conducted an empirical study on 659 subjects to uncover the determinants of the performance of bodyweight exercise behavior. After modeling 2 popular bodyweight exercise behaviors using a virtual coach, they showed that perceived self-efficacy and perceived social support are the strongest determinants of this exercise behavior.

In a cohort of 10 patients with prediabetes, Srivastava et al [52] recently offered a module consisting of a Web app supporting diabetes prevention education and a mobile app with an electronic diary and virtual coach. Thanks to an efficient review of user performance and the ability to send support notifications from the user's coach or physician, the module reached a high success rate (60%), allowing patients at high risk for diabetes to be engaged in the process of lifestyle modification.

Chronic Diseases

VC could play a fundamental role for people with chronic diseases, providing them with support for the self-management of their problems at home. This explains why some authors have begun to evaluate the use of e-coaches in communities of no longer healthy subjects. In 2008, Allen et al [17] developed an internet-based health coaching intervention to enhance patient-provider communication regarding 3 common conditions: chronic pain, depression, and impaired mobility. This efficient and low-cost approach offered an innovative opportunity to improve patient and clinician partnerships in managing chronic conditions. An internet-based health coaching intervention that can offer a significant benefit to many patients, but differs substantially from usual nursing care, was also implemented to empower patients toward collaborative self-management. Providing and evaluating the experience gathered within a community of patients with Crohn disease, Kreps et al [39] showed that VCs and comfortable human-computer interfaces (based on user-centered AI approaches) could stimulate active information processing and adoption of new thoughts, such as motivation and behavior changes. A proof of concept of a pervasive gamified platform framework that is a combination of integrated sensors (wearables), a VC, and serious games in health care has been recently reported by Klaassen et al [47]. Interestingly, they combined different technology components to promote positive health behaviors among young people with type 1 diabetes.

Ritchie et al [28] conducted a study on the use of the system and readmission outcomes of a pragmatic randomized trial of e-coach (as the VC was called in this paper). In the study, the system was used in a diverse sample of complex medical patients with congestive heart failure and COPD from a wide geographic region in the Southern United States. Although 30-day rehospitalization rates did not statistically differ between the e-coach and usual postdischarge care groups, in the COPD subgroup the e-coach was associated with significantly fewer days in the hospital. This indicated that interventions may need to be disease-specific to decrease rehospitalization rates and to ensure adequate postdischarge care. This study addressed the problem of how to manage care transition from the hospital to the home for people with two different critical illnesses. E-coach, supported by an interactive voice response (IVR) system, did not reduce the rehospitalization of these subjects. This evidence led the authors to conclude that, to date, a combination of an IVR and personal contact with the treating physician is the solution to that has the most beneficial effect on reducing preventable readmissions and providing optimal transitional support for these complex patients. In the cardiological field, Adams et al [42] described a single case of a patient who was a powerlifter and returned to his sport after coronary artery bypass grafting with long-distance coaching by the cardiac rehabilitation staff. Through high-intensity training that was complemented by phone and email support, he lifted heavier loads than he had before the bypass grafting.

Regarding the neurological diseases, Ellis et al [37] conducted a phase 1, single-group, nonrandomized clinical trial in Parkinson disease (PD). The intervention was designed to increase the physical activity of 20 subjects with mild-to-moderate PD by promoting additional daily walking using a pedometer and brief daily interactions with the computer-animated virtual exercise coach. Patients successfully used the virtual coach at home over a 1-month period, demonstrating the initial feasibility of this approach.

With a virtual therapist designed for speech therapy that works across devices, and built-in Web monitoring, scheduling, and communication technologies, van Vuuren et al [41] recruited 8 participants with chronic aphasia to receive intensive computer-based script training differing in the amount of high or low cuing provided during treatment. With the limitations of a small sample size and the lack of separation of auditory and visual cuing, the overall effect size for the computerized treatment was large and like what would be expected when treatment was delivered by a speech-language pathologist.

With the role of virtual nursing avatars (VNAs) considered to be promising and evolving [66] in the context of chronic disease, Cotè et al [35,43,59,61] promoted the application of VNAs

through a Web app to support the management of antiretroviral therapy and the adoption of healthy behaviors in HIV patients.

In a 3-month pre- to postdesign pilot study, the promotion of HIV medication adherence was also explored in 43 patients using a theory-based, mobile-delivered, ECA to provide information and behavioral skills. They showed good acceptability and preliminary efficacy in improving drug adherence [51].

The promotion of therapy adherence in other chronic diseases was also described through a Web-based nursing intervention in some pilot RCTs [19,31], whose results support the feasibility and acceptability of the virtual nursing sessions in a small sample of patients.

Psychology

The use of VC has also involved the field of, both in offering new educational approaches or psychotherapeutic support, and in trying to understand the relationship between ECA and user. In 2013, Hudlicka [38] reported that the virtual coach–based training of mindfulness was potentially more effective than a self-administered program, including written and audio material, in a group of students initiated to this practice. Although it had a low number of participants, the study highlighted the possible effectiveness of virtual training and coaching in ending or acquiring a behavior that may be deleterious (eg, smoking) or healthy (eg, exercise) in people exposed to medical risk. A virtual agent was also used with good results in favoring adherence to a therapeutic task within a psychoeducational process [45]. However, the study showed that if psychoeducation was presented in the text, it resulted in better adherence than if the agent offered it verbally. Indeed, we know how important the mimicry and the expressiveness of the ECA are for its effective interaction with the user. These 2 aspects have been addressed in 2 interesting experiments reported in a recent article by Stevens et al [44] (see “Interaction between virtual coach and the user”). Recently, ECA applications in clinical psychology have been reviewed but the authors concluded that these applications are still limited [55].

Risk Factors Reduction

ECAs and VC services with an integrated cognitive training system, interactive clinical scales and tests to identify chewing and swallowing difficulties, with integrated pedometers and other wearable sensors (fall, freezing, or motor fluctuation detection systems), as well as a home-based interactive information environment can be used to identify and prevent disease-related complications at home or outdoors. This type of coaching system has been demonstrated in health problems such as obesity and a sedentary lifestyle [20,22,27]. This user interface would also provide daily motor, cognitive, and behavioral exercises and tips, monitor lifestyle and take care of treatment adherence monitoring, with AI to solve basic inquiries, and would provide contact with urgent or nonurgent health care responses if needed. The hardware of this system would be a combination of home-based wall microphones, speakers, and screens and a mobile phone–based system. A comparable system for heart failure was implemented for the My Heart IST-2002-507816 project [33]. This system was designed to be remotely supervised by humans specialized in health care, such as nurses, providing an environment for transitional care of patients. This would have had an additional positive impact on the prevention of hospital admission–related complications and in reducing sanitary burden and costs by minimizing admission days in the hospital and out of community [21].

Treatment Adherence

Any implemented VC service such as ECA should also take specific care of treatment adherence monitoring. In this sense, the creation of a virtual social environment may include that patients could exchange information with other patients or could directly contact their reference health care professionals for treatment-related inquiries [27]. However, the adherence to the treatment or intervention very much depends on the humanity of the ECA, such as showing empathy, emotions, or using human-like speech. Those devices that are monitored and controlled remotely by humans have been shown to be particularly effective [30,46]. Kreps and Neuhauser demonstrated that it is possible to provide computer-mediated support to patients to track their medication adherence [39], and this evidence may be very useful in preventing motor fluctuations and dyskinesias in PD patients, where the chronology of treatment is important [37].

Discussion

Findings

This scoping review shows that in the last 10 years, there has been a growing interest in implementing VC programs aimed at stimulating and guiding users toward positive behaviors in the various sectors of the health care world. In the review, we evaluated all available evidence on the use of VC, defined a coaching program, and suggested problem-solving strategies helping users to face specific medical conditions. In most cases, a successful VC approach must consist of participant-involving methods, with a focus on the individual patient’s lifestyle conditions. In addition, the review allowed us to map the key concepts currently tied to the users’ or patients’ VC, such as function and symptoms, medication, dietary recommendations, recognition and coping with alarm signals of worsening of the disease, psychological and emotional reactions in connection with the disease, and exercise strategies. The patients’ relatives or other acquaintances can be involved depending on the patients’ needs and requests, and education must be conducted in a combined individual and group-based setting.

Recently, another review examined eHealth interventions with the combined use of self-tracking and persuasive e-Coaching [54]. It is interesting to note that the review provides information on which key components can be effective on health outcomes, usability, and adherence. Although the analysis is relevant for future strategies that combine self-tracking and persuasive e-Coaching, it addresses the topic of the virtual coach differently from our review, which aimed to analyze the context and the medical implications of the interventions carried out rather than the aspects that can influence the effective use of technological solutions. We extended the investigation to the study protocols, the types of interface adopted, and the medical areas of interest, adding a dedicated section for the latter, and we decided to

gather this information as it would allow us to assess the state of the VC application related to its degree of use and diffusion in the medical health field.

General Implications in the Medical Sector

Regarding the medical care context, to date, the use of virtual coach technology–assisted care seems to be more represented but it has mainly been implemented for healthy subjects with the aim of improving their health behaviors and reducing the risk factors for cardio- and cerebrovascular diseases, such as obesity or physical inactivity. On the contrary, only 7 studies considered groups of patients with chronic diseases (ie, depression, heart failure, COPD, PD, Crohn disease, diabetes, and urinary incontinence) with the intent of demonstrating the usefulness of different internet-based health coaching interventions. These were aimed, for example, at improving communication with nursing staff, reducing the rate of rehospitalization of patients, or promoting their daily physical activity.

It is noteworthy that this review found 16 RCT studies among the 46 articles selected for concrete coaching interventions in the medical health field in the last 10 years. However, although this may represent an interesting fact for a topic that has appeared over the last 15 years, it deserves some specific considerations. First, as the majority of RCT studies involved healthy subjects, although often constitutionally overweight or in old age, their primary aim was often to promote healthy behavior. Moreover, with particular reference to this type of intervention on a healthy population, RCT studies have so far preferentially evaluated the use of VC for the purposes of primary prevention [18,21,22,24-27,29,30,32] and only in 4 cases [17,19,28,31] have they speculated about its application in chronic pathological conditions. Thus, most RCT studies have, above all, demonstrated how various technologies used to activate VC are feasible and acceptable and, consequently, how they can have a potential impact on preserving public health.

In relation to the ambitious goal of helping with the implementation of primary prevention, it is obvious that both the limited number of subjects studied and the short time of use of the technological solutions have not yet allowed research to sustain the absolute effectiveness of VC in producing a change in harmful behavior or maintaining good habits that can positively affect people's health. Conversely, it is known that studies related to lifestyles and how these can be modified in a positive and healthy way require several years of monitoring of a large population. Second, although 6 RCT studies evaluating physical activity and 3 being overweight are present, the comparability of results within each of the two groups is inadequate for differences in average age and sample size of the studied populations, in the types of interface, and in the used outcome measures.

Most of the remaining studies, definable as exploratory, return little information beyond the interest of some authors in evaluating the application of VC in the management of some pathological conditions.

Health promotion can be achieved by using ECA or IVR systems, as the usability and acceptance of these devices is usually high [24,30,46]. They have been used for engaging the elderly in regular physical exercise [24,30,46], and they have been applied in combination with tracking devices and sensors, such as pedometers, to promote a healthy lifestyle in older adults with obesity and a sedentary lifestyle, [20,22,27] and to enhance an adult's social interaction.

The therapeutic alliance between the patient and the health professional has also been considered an integrative element in all forms of therapy, regarding both the efficacy of treatment and the improvement of adherence. Interestingly, ECAs, especially those exhibiting empathic behaviors, can potentially deliver nonjudgmental support like real face to-face communication, and thus may compensate for clinical failures in some groups of patients. ECA might also be used to promote healthy eating habits specific to neurologic and cardiologic patients. In addition, taking care of nutrition in patients is very important, particularly during pharmacological treatments as many nutrients may interact with drugs (eg, levodopa), thus interfering with their intestinal absorption. Although the ideal characteristics of an ECA developing an empathic relationship with the user are still under study, this interface seems to bring good acceptability in many studies and could be the best solution today in terms of an interface for controlling the regular intake of drugs in patients with chronic diseases.

Cardiological and Neurological Areas

The review has revealed a certain interest in cardiological and neurological diseases which, as is well known, often require changes in lifestyle, adherence to drug therapies, and specific rehabilitation programs. Specifically, cardiological and neurological diseases are especially relevant for people as they age. They include cardiovascular diseases, such as heart attack and heart failure, cerebrovascular diseases, such as stroke, or neurodegenerative disorders such as PD. Admittedly, these diseases have quite different pathogenic mechanisms and therapeutic treatment, with an acute onset or a chronic course. However, the structure of the rehabilitation programs is quite similar, even if the care plans and their basis might differ at some points. Their main objective is to help the patients in improving their quality of life as much as possible through recovery of their fitness levels and motor and cognitive disabilities. People should be supported in relearning or regaining the ability to deal with their home environment without (extensive) help. Thanks to the progress in acute treatment (eg, coronary angioplasty, thrombolysis, the concept of stroke units, and dysphagia management), most people survive acute pathological events but they are often affected by long-term disability.

Consequently, rehabilitation (mainly as organized inpatient multidisciplinary rehabilitation) remains one of the cornerstones of disability treatment. Rehabilitation strategies help survivors to relearn skills that are lost. In the literature, there are models [67,68] that illustrate disease progression with and without rehabilitation, emphasizing its importance in improving functional performance. The main idea is that rehabilitation is necessary to develop and expand functional recovery in acute

disease (eg, heart attack or stroke) and to slow down cognitive and motor function decline in chronic disease (eg, heart failure or PD). Currently, the adjustment of chronic drug therapy and rehabilitation therapy may start from the acute care hospital after the person's overall condition has been stabilized, and it should be a long ongoing process continued after being out of clinic for months or years. However, the surveillance of correct drug intake and the continuation of rehabilitation treatment are difficult in the domestic environment, especially if not supported adequately by informal caregivers and in the presence of frailty conditions likely to be relevant for the older population. So, in most cases, the correct drug intake may not occur and the rehabilitation process will end following clinical discharge, so patients will have a risk of a clinically unstable condition or a gradual increase in their disability level by detraining. In this scenario, the intervention of a VC would have significant and positive implications in terms of home therapeutic management of patients suffering from these diseases.

Lack of Continuity of Care

The lack of continuity of care leads to a gap between the real patient's condition and the level of recovery that they could reach with a constant, well-structured rehabilitation process [51,52]. To maximize the rehabilitation outcome, the patient's motivation should be sustained during exercise repetition. Patients need to be able to practice motor actions in different tasks and environmental contexts to develop motor schemata that are versatile enough to meet the situations they encounter in daily life. The difficulty level of the motor and cognitive task assigned to the patient, together with their subjective awareness of the obtained global performance and the quantity and quality of feedback received during training, can influence patient motivation and produce different means of acting and different performances.

In addition to improving motor and cognitive manifestations by proper monitoring and rehabilitation programs, one of the main elements determining quality of life is related to rehabilitation strategies oriented to promote independence in basic and instrumental activities of daily living. For all these reasons, it is time to design and implement coaching systems capable of activating the appropriate strategies to improve the motor and cognitive disabilities of patients, including those promoting independence in personal hygiene, dressing, self-feeding, using the telephone or other forms of communication, and taking prescribed medications. Thus, continuously applied and personalized rehabilitation programs help people with disabilities to improve the body functions that limit reinsertion into their home or community, living independently, and participating in education, the labor market, and civic life [41]. Currently, the major disadvantages of home-based rehabilitation solutions are the lack of specialized equipment and insufficient alignment and adaptability of the technical possibilities to the individual care needs and abilities, but *ad hoc* VC systems could be the answer to these problems.

Future Perspective of Virtual Coaching System

The review of the VC literature confirms an increased diffusion of technologies with strong interaction with human beings in the world of health and medicine. The increase in interest on this issue is evident not only in the number of publications since the beginning of the current year but also in the guiding vision of the EU's Horizon 2020 research framework program.

If up to now most studies have tested the goodness of the technological interlocutor and the availability of the human one, in the future, we will try to strengthen this relationship in the medical sector as we are working to foster patients' engagement toward a process of self-awareness and self-confidence in health management. Nevertheless, advanced modes of presence will lead to ubiquity of the VC and allow further interaction modes. Integrating the VC within the real world in terms of an augmented reality (eg, the avatar could also be present in the car) will broaden its abilities to guide the user and therefore increase the literacy of the suggested actions. Advanced sensing and acting levels allow seamless adaptation in relation to the people's conditions. Prospectively, multiple linked coaches will allow better monitoring and guidance of the user. Furthermore, the to-be virtual coach will not only improve basic treatment scenarios, but the advanced AI may also infer new treatment strategies.

Hopefully, these changes will allow us to deal with increasingly complex clinical situations, mainly belonging to the world of chronic diseases and disabilities. The lesson acquired so far, aimed at promoting coaching interventions often on healthy subjects, independent thinking, and acting, represents an important step in transferring the experience of the interventions on the major themes of primary prevention to the more complicated and specific ones of secondary prevention.

Thus, the future of VC for frail or sick people is expected to provide solid evidence about the usefulness and effectiveness in preserving physical, cognitive, mental, and social well-being for as long as possible. More specifically, the VC will educate and finally empower patients to increase their adherence to drug therapy and to pursue physical and cognitive rehabilitation programs to regain independence with a healthier lifestyle and a better quality of life.

Strengths and Limitations of This Scoping Review

This scoping review was guided by a protocol reviewed by a research team that adopted Arksey and O'Malley's definition for scoping reviews at the outset of the study and found that their simple definition was generally useful in guiding study selection. To ensure a broad search of the literature, the search strategy included 3 electronic bibliographic databases, the reference list of different articles, the websites of relevant journals, and the snowball technique. Each citation and article were then reviewed by 3 independent reviewers belonging to different disciplines (medicine and bioengineering) who met at regular intervals to resolve conflicts.

For several reasons, this review brings new information to the reader's attention. First, it revised the definition of virtual coach, demonstrating the current validity of the first description made by Ding and how it can still include the technological developments of recent years aimed at improving interaction between virtual coach and user. Second, with an innovative approach, the review significantly aimed at highlighting which medical care sectors were considered to apply VC. Considering

this extensive analysis, the review aimed to return useful information to bioengineers regarding the current potentialities of the medical application of VC and look at its possible future uses.

Nevertheless, some limitations are worth noting. First, because of variability in the conduct of scoping reviews, we are aware that there is a need for methodological standardization to ensure the strength of evidence. Second, this review may not have identified all the papers in the published literature despite attempts to be as comprehensive as possible. Our search algorithm included 9 different terms (and their declination) previously used to describe the VC devices and services; however, other terms may also exist. Although our search included 3 databases (ie, PubMed, Scopus, and Embase), the overall search strategy may have been biased toward health and sciences. Moreover, we may have missed some papers in the gray literature.

Conclusions

The continuous technological progress is changing the impact on users of the use of a humanoid avatar, which is an artificial figure apparently substitutive for a health professional. This relationship could entail numerous preconceptions, such as the lack of a direct dialogue between doctor and patient. As already underlined, a trusty relationship will be based on the good interaction between the user and the avatar and on how much the latter is recognized as an element of a high-quality medical care process. However, the psychological implications of this relationship can only be known through the concrete application of these technologies. Based on our analysis, we must look forward with confidence, aware that there are still several open questions. In a broader vision, many articles on VC have directly or indirectly addressed medical aspects of utmost importance in both primary (eg, risk factors reduction) and secondary (eg, treatment adherence) prevention.

However, the review gained evidence that VC is more rarely applied in medical care for secondary prevention. As a matter of fact, although physical activity is mostly encouraged by a VC system, we know that it cannot be compared with a real rehabilitation program, whose indication is of extreme importance in patients suffering from acute disabling events or motor functional impairments because of several chronic conditions. Therefore, in the current scenario of VC application, rehabilitation is the great absentee. Thus, looking at a secondary prevention of chronic diseases, VC has a concrete potential to improve medication adherence and to extend existing rehabilitation programs for patients, providing a great opportunity to guarantee the continuity of care between hospital and home. Considering these considerations, it is essential that the constructive dialogue between bioengineers and clinicians in creating innovative technological platforms continues and improves to deal synergistically with new challenges in the world of medical care.

Acknowledgments

This work was partially supported by the EU within the vCare project (funded by Horizon 2020 research and innovation program under grant agreement No 769807). The authors would like to acknowledge all vCare colleagues and the Need Institute Foundation for their supportive collaboration.

Authors' Contributions

PT, HS, IS, KG, MCa, and MCo conceived of the presented idea. PT, IS, and MCo analyzed the data and interpreted the results. HS, EJ, KG, MCa, IG, JCG, SB, CS, SK, and SA aided in interpreting the results. PT and MCo took the lead in writing the manuscript. All authors provided critical feedback and helped shape the manuscript.

Conflicts of Interest

None declared.

References

1. European Heart Network. 2017. European Cardiovascular Disease Statistics 2017 URL: http://www.ehnheart.org/cvd-statistics.html [accessed 2018-10-29] [WebCite Cache ID 73WrO6fbd]
2. European Commission. 2017. Eu Labour Force Survey - Ad Hoc Modules URL: https://ec.europa.eu/eurostat/statistics-explained/index.php/EU_labour_force_survey_-_ad_hoc_modules [accessed 2018-10-29] [WebCite Cache ID 73WragS6G]
3. Kolominsky-Rabas PL, Heuschmann PU, Marschall D, Emmert M, Baltzer N, Neundörfer B, et al. Lifetime cost of ischemic stroke in Germany: results and national projections from a population-based stroke registry: the Erlangen stroke project. Stroke 2006 May;37(5):1179-1183. [doi: 10.1161/01.STR.0000217450.21310.90] [Medline: 16574918]
4. Parker SG, Oliver P, Pennington M, Bond J, Jagger C, Enderby PM, et al. Rehabilitation of older patients: day hospital compared with rehabilitation at home. A randomised controlled trial. Health Technol Assess 2009 Aug;13(39):1-143, iii [FREE Full text] [doi: 10.3310/hta13390] [Medline: 19712593]
5. Ries AL, Bauldoff GS, Carlin BW, Casaburi R, Emery CF, Mahler DA, et al. Pulmonary rehabilitation: joint ACCP/AACVPR evidence-based clinical practice guidelines. Chest 2007 May;131(5 Suppl):4S-42S. [doi: 10.1378/chest.06-2418] [Medline: 17494825]

6. Gabrielli S, Dianti M, Maimone R, Betta M, Filippi L, Ghezzi M, et al. Design of a mobile app for nutrition education (TreC-LifeStyle) and formative evaluation with families of overweight children. JMIR Mhealth Uhealth 2017 Apr 13;5(4):e48 [FREE Full text] [doi: 10.2196/mhealth.7080] [Medline: 28408361]
7. Johnson WL, LaBore C, Chiu YC. A Pedagogical Agent for Psychosocial Intervention on a Handheld Computer. In: Proceedings of the AAAI Fall Symposium on Dialogue Systems for Health Communication. 2004 Presented at: AAAI'04; October 21-24, 2004; Arlington, Virginia p. 22-24.
8. van den Berg MH, Schoones JW, Vlieland TP. Internet-based physical activity interventions: a systematic review of the literature. J Med Internet Res 2007 Sep 30;9(3):e26 [FREE Full text] [doi: 10.2196/jmir.9.3.e26] [Medline: 17942388]
9. Bickmore T, Gruber A, Picard R. Establishing the computer-patient working alliance in automated health behavior change interventions. Patient Educ Couns 2005 Oct;59(1):21-30. [doi: 10.1016/j.pec.2004.09.008] [Medline: 16198215]
10. European Commission. 2007. Personalised Coaching for Well-being and Care of People as They Age URL: http://ec.europa.eu/research/participants/portal/desktop/en/opportunities/h2020/topics/sc1-pm-15-2017.html [accessed 2019-07-25] [WebCite Cache ID 73WrpdVad]
11. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodol 2005 Feb;8(1):19-32. [doi: 10.1080/1364557032000119616]
12. Mays N, Roberts E, Popay J. Synthesising research evidence. In: Fulop N, Allen P, Black N, Clarke A, editors. Studying the Organisation and Delivery of Health Services: Research Methods. London: Routledge; 2001:188-220.
13. National Center for Biotechnology Information. PubMed URL: http://www.ncbi.nlm.nih.gov/pubmed/ [accessed 2019-07-25] [WebCite Cache ID 73Ws4LW9J]
14. Scopus. URL: http://www.scopus.com [accessed 2019-07-25] [WebCite Cache ID 73Ws8bIkL]
15. Embase. URL: http://www.embase.com [accessed 2019-07-25] [WebCite Cache ID 73Ws8bIkL]
16. Mendeley. URL: http://www.mendeley.com [accessed 2019-07-25] [WebCite Cache ID 73WsB8ZsB]
17. Allen M, Iezzoni LI, Huang A, Huang L, Leveille SG. Improving patient-clinician communication about chronic conditions: description of an internet-based nurse e-coach intervention. Nurs Res 2008;57(2):107-112. [doi: 10.1097/01.NNR.0000313478.47379.98] [Medline: 18347482]
18. Gabriele JM, Carpenter BD, Tate DF, Fisher EB. Directive and nondirective e-coach support for weight loss in overweight adults. Ann Behav Med 2011 Apr;41(2):252-263 [FREE Full text] [doi: 10.1007/s12160-010-9240-2] [Medline: 21108032]
19. Martorella G, Côté J, Racine M, Choinière M. Web-based nursing intervention for self-management of pain after cardiac surgery: pilot randomized controlled trial. J Med Internet Res 2012 Dec 14;14(6):e177 [FREE Full text] [doi: 10.2196/jmir.2070] [Medline: 23241361]
20. Watson A, Bickmore T, Cange A, Kulshreshtha A, Kvedar J. An internet-based virtual coach to promote physical activity adherence in overweight adults: randomized controlled trial. J Med Internet Res 2012 Jan 26;14(1):e1 [FREE Full text] [doi: 10.2196/jmir.1629] [Medline: 22281837]
21. Bickmore TW, Schulman D, Sidner C. Automated interventions for multiple health behaviors using conversational agents. Patient Educ Couns 2013 Aug;92(2):142-148 [FREE Full text] [doi: 10.1016/j.pec.2013.05.011] [Medline: 23763983]
22. Bickmore TW, Silliman RA, Nelson K, Cheng DM, Winter M, Henault L, et al. A randomized controlled trial of an automated exercise coach for older adults. J Am Geriatr Soc 2013 Oct;61(10):1676-1683. [doi: 10.1111/jgs.12449] [Medline: 24001030]
23. Wijsman CA, Westendorp RG, Verhagen EA, Catt M, Slagboom PE, de Craen AJ, et al. Effects of a web-based intervention on physical activity and metabolism in older adults: randomized controlled trial. J Med Internet Res 2013 Nov 6;15(11):e233 [FREE Full text] [doi: 10.2196/jmir.2843] [Medline: 24195965]
24. Friederichs S, Bolman C, Oenema A, Guyaux J, Lechner L. Motivational interviewing in a web-based physical activity intervention with an avatar: randomized controlled trial. J Med Internet Res 2014 Feb 13;16(2):e48 [FREE Full text] [doi: 10.2196/jmir.2974] [Medline: 24550153]
25. Vroege DP, Wijsman CA, Broekhuizen K, de Craen AJ, van Heemst D, van der Ouderaa FJ, et al. Dose-response effects of a web-based physical activity program on body composition and metabolic health in inactive older adults: additional analyses of a randomized controlled trial. J Med Internet Res 2014 Dec 4;16(12):e265 [FREE Full text] [doi: 10.2196/jmir.3643] [Medline: 25486673]
26. Broekhuizen K, de Gelder J, Wijsman CA, Wijsman LW, Westendorp RG, Verhagen E, et al. An internet-based physical activity intervention to improve quality of life of inactive older adults: a randomized controlled trial. J Med Internet Res 2016 Apr 27;18(4):e74 [FREE Full text] [doi: 10.2196/jmir.4335] [Medline: 27122359]
27. Leahey TM, Fava JL, Seiden A, Fernandes D, Doyle C, Kent K, et al. A randomized controlled trial testing an internet delivered cost-benefit approach to weight loss maintenance. Prev Med 2016 Nov;92:51-57 [FREE Full text] [doi: 10.1016/j.ypmed.2016.04.013] [Medline: 27095323]
28. Ritchie CS, Houston TK, Richman JS, Sobko HJ, Berner ES, Taylor BB, et al. The e-coach technology-assisted care transition system: a pragmatic randomized trial. Transl Behav Med 2016 Sep;6(3):428-437 [FREE Full text] [doi: 10.1007/s13142-016-0422-8] [Medline: 27339715]

29. Gardiner PM, McCue KD, Negash LM, Cheng T, White LF, Yinusa-Nyahkoon L, et al. Engaging women with an embodied conversational agent to deliver mindfulness and lifestyle recommendations: a feasibility randomized control trial. Patient Educ Couns 2017 Sep;100(9):1720-1729 [FREE Full text] [doi: 10.1016/j.pec.2017.04.015] [Medline: 28495391]
30. King AC, Campero I, Sheats JL, Sweet CM, Garcia D, Chazaro A, et al. Testing the comparative effects of physical activity advice by humans vs computers in underserved populations: the COMPASS trial design, methods, and baseline characteristics. Contemp Clin Trials 2017 Oct;61:115-125 [FREE Full text] [doi: 10.1016/j.cct.2017.07.020] [Medline: 28739541]
31. Côté J, Fortin MC, Auger P, Rouleau G, Dubois S, Boudreau N, et al. Web-based tailored intervention to support optimal medication adherence among kidney transplant recipients: pilot parallel-group randomized controlled trial. JMIR Form Res 2018 Jul 19;2(2):e14 [FREE Full text] [doi: 10.2196/formative.9707] [Medline: 30684400]
32. Hui SS, Xie YJ, Kwok RC, Tam EW, Mak WW, Mo PK. Follow your virtual trainer (FYVT): a randomised controlled trial protocol of IT-based lifestyle intervention programme to promote physical activity and health among middle-aged Hong Kong Chinese. BMJ Open 2018 Feb 3;8(2):e017908 [FREE Full text] [doi: 10.1136/bmjopen-2017-017908] [Medline: 29431125]
33. Guillen SG, Traver V, Sala P, Traver M, Romagnoli M, Arredondo MT, et al. Interactive Information Environment for the Controlled Practice of Physical Training to Improve Cardiovascular Fitness. In: Proceedings of the IEEE Engineering in Medicine and Biology 27th Annual Conference. 2006 Presented at: EMBC'06; January 18-18, 2006; Shanghai, China p. 2211-2214. [doi: 10.1109/IEMBS.2005.1616902]
34. Segerståhl K, Oinas-Kukkonen H. Designing personal exercise monitoring employing multiple modes of delivery: implications from a qualitative study on heart rate monitoring. Int J Med Inform 2011 Dec;80(12):e203-e213. [doi: 10.1016/j.ijmedinf.2011.08.011] [Medline: 21963231]
35. Côté J, Rouleau G, Godin G, Ramirez-Garcìa P, Guéhéneuc YG, Nahas G, et al. Acceptability and feasibility of a virtual intervention to help people living with HIV manage their daily therapies. J Telemed Telecare 2012 Oct;18(7):409-412. [doi: 10.1258/jtt.2012.120218] [Medline: 23034932]
36. Novielli N, Mazzotta I, de Carolis B, Pizzutilo S. Analysing user's reactions in advice-giving dialogues with a socially intelligent ECA. Cogn Process 2012 Oct;13(Suppl 2):487-497. [doi: 10.1007/s10339-011-0420-1] [Medline: 21997309]
37. Ellis T, Latham NK, DeAngelis TR, Thomas CA, Saint-Hilaire M, Bickmore TW. Feasibility of a virtual exercise coach to promote walking in community-dwelling persons with Parkinson disease. Am J Phys Med Rehabil 2013 Jun;92(6):472-81; quiz 482 [FREE Full text] [doi: 10.1097/PHM.0b013e31828cd466] [Medline: 23552335]
38. Hudlicka E. Virtual training and coaching of health behavior: example from mindfulness meditation training. Patient Educ Couns 2013 Aug;92(2):160-166 [FREE Full text] [doi: 10.1016/j.pec.2013.05.007] [Medline: 23809167]
39. Kreps GL, Neuhauser L. Artificial intelligence and immediacy: designing health communication to personally engage consumers and providers. Patient Educ Couns 2013 Aug;92(2):205-210. [doi: 10.1016/j.pec.2013.04.014] [Medline: 23683341]
40. Silveira P, van de Langenberg R, van Het RE, Daniel F, Casati F, de Bruin ED. Tablet-based strength-balance training to motivate and improve adherence to exercise in independently living older people: a phase II preclinical exploratory trial. J Med Internet Res 2013 Aug 12;15(8):e159 [FREE Full text] [doi: 10.2196/jmir.2579] [Medline: 23939401]
41. van Vuuren S, Cherney LR. A virtual therapist for speech and language therapy. Intell Virtual Agents 2014 Jan 1;8637:438-448 [FREE Full text] [doi: 10.1007/978-3-319-09767-1_55] [Medline: 25938137]
42. Adams R, Adams J, Qin H, Bilbrey T, Schussler JM. Virtual coaching for the high-intensity training of a powerlifter following coronary artery bypass grafting. Proc (Bayl Univ Med Cent) 2015 Jan;28(1):75-77 [FREE Full text] [doi: 10.1080/08998280.2015.11929196] [Medline: 25552808]
43. Côté J, Rouleau G, Ramirez-Garcia P, Bourbonnais A. Virtual nursing intervention adjunctive to conventional care: the experience of persons living with HIV. JMIR Res Protoc 2015 Oct 20;4(4):e124 [FREE Full text] [doi: 10.2196/resprot.4158] [Medline: 26487327]
44. Stevens CJ, Pinchbeck B, Lewis T, Luerssen M, Pfitzner D, Powers DM, et al. Mimicry and expressiveness of an ECA in human-agent interaction: familiarity breeds content!. Comput Cogn Sci 2016;2(1):1 [FREE Full text] [doi: 10.1186/s40469-016-0008-2] [Medline: 27980890]
45. Tielman ML, Neerincx MA, van Meggelen M, Franken I, Brinkman WP. How should a virtual agent present psychoeducation? Influence of verbal and textual presentation on adherence. Technol Health Care 2017 Dec 4;25(6):1081-1096 [FREE Full text] [doi: 10.3233/THC-170899] [Medline: 28800346]
46. Chi NC, Sparks O, Lin SY, Lazar A, Thompson HJ, Demiris G. Pilot testing a digital pet avatar for older adults. Geriatr Nurs 2017;38(6):542-547 [FREE Full text] [doi: 10.1016/j.gerinurse.2017.04.002] [Medline: 28479082]
47. Klaassen R, Bul KC, op den Akker R, van der Burg GJ, Kato PM, di Bitonto P. Design and evaluation of a pervasive coaching and gamification platform for young diabetes patients. Sensors (Basel) 2018 Jan 30;18(2):E402 [FREE Full text] [doi: 10.3390/s18020402] [Medline: 29385750]
48. Oyibo K, Adaji I, Vassileva J. Social cognitive determinants of exercise behavior in the context of behavior modeling: a mixed method approach. Digit Health 2018;4:2055207618811555 [FREE Full text] [doi: 10.1177/2055207618811555] [Medline: 30479828]

49. Richards D, Caldwell P. Improving health outcomes sooner rather than later via an interactive website and virtual specialist. IEEE J Biomed Health Inform 2018 Sep;22(5):1699-1706. [doi: 10.1109/JBHI.2017.2782210] [Medline: 29990229]
50. Suganuma S, Sakamoto D, Shimoyama H. An embodied conversational agent for unguided internet-based cognitive behavior therapy in preventative mental health: feasibility and acceptability pilot trial. JMIR Ment Health 2018 Jul 31;5(3):e10454 [FREE Full text] [doi: 10.2196/10454] [Medline: 30064969]
51. Dworkin MS, Lee S, Chakraborty A, Monahan C, Hightow-Weidman L, Garofalo R, et al. Acceptability, feasibility, and preliminary efficacy of a theory-based relational embodied conversational agent mobile phone intervention to promote HIV medication adherence in young HIV-positive African American MSM. AIDS Educ Prev 2019 Feb;31(1):17-37. [doi: 10.1521/aeap.2019.31.1.17] [Medline: 30742481]
52. Srivastava P, Verma A, Geronimo C, Button TM. Behavior stages of a physician- and coach-supported cloud-based diabetes prevention program for people with prediabetes. Sage Open Med 2019;7:2050312119841986 [FREE Full text] [doi: 10.1177/2050312119841986] [Medline: 31105938]
53. Ding D, Liu HY, Cooper R, Cooper RA, Smailagic A, Siewiorek D. Virtual coach technology for supporting self-care. Phys Med Rehabil Clin N Am 2010 Feb;21(1):179-194. [doi: 10.1016/j.pmr.2009.07.012] [Medline: 19951785]
54. Lentferink AJ, Oldenhuis HK, de Groot M, Polstra L, Velthuijsen H, van Gemert-Pijnen JE. Key components in ehealth interventions combining self-tracking and persuasive ecoaching to promote a healthier lifestyle: a scoping review. J Med Internet Res 2017 Aug 1;19(8):e277 [FREE Full text] [doi: 10.2196/jmir.7288] [Medline: 28765103]
55. Provoost S, Lau HM, Ruwaard J, Riper H. Embodied conversational agents in clinical psychology: a scoping review. J Med Internet Res 2017 May 9;19(5):e151 [FREE Full text] [doi: 10.2196/jmir.6553] [Medline: 28487267]
56. Vaidyam AN, Wisniewski H, Halamka JD, Kashavan MS, Torous JB. Chatbots and conversational agents in mental health: a review of the psychiatric landscape. Can J Psychiatry 2019 Jul;64(7):456-464. [doi: 10.1177/0706743719828977] [Medline: 30897957]
57. Salvemini A, D'Onofrio G, Ciccone F, Greco A, Tullio A, Addante F, et al. Insomnia and information and communication technologies (ICT) in elderly people: a systematic review. Med Sci (Basel) 2019 Jun 15;7(6):E70 [FREE Full text] [doi: 10.3390/medsci7060070] [Medline: 31208061]
58. de Rosis F, Novielli N, Carofiglio V, Cavalluzzi A, de Carolis B. User modeling and adaptation in health promotion dialogs with an animated character. J Biomed Inform 2006 Oct;39(5):514-531 [FREE Full text] [doi: 10.1016/j.jbi.2006.01.001] [Medline: 16524784]
59. Côté J, Ramirez-Garcia P, Rouleau G, Saulnier D, Guéhéneuc YG, Hernandez A, et al. A nursing virtual intervention: real-time support for managing antiretroviral therapy. Comput Inform Nurs 2011;29(1):43-51. [doi: 10.1097/NCN.0b013e3181f9dc02] [Medline: 21099544]
60. Pérez J, Cerezo E, Serón FJ, Rodríguez LF. A cognitive-affective architecture for ECAs. Biol Inspir Cogn Arc 2016 Oct;18:33-40. [doi: 10.1016/j.bica.2016.10.002]
61. Côté J, Cossette S, Ramirez-Garcia P, Rouleau G, Auger P, Boudreau F, et al. Improving health and reducing comorbidity associated with HIV: the development of TAVIE en sante, a web-based tailored intervention to support the adoption of health promoting behaviors among people living with HIV. Biomed Res Int 2017;2017:4092304 [FREE Full text] [doi: 10.1155/2017/4092304] [Medline: 28393077]
62. Fadhil A, Wang Y, Reiterer H. Assistive conversational agent for health coaching: a validation study. Methods Inf Med 2019 Jun;58(1):9-23. [doi: 10.1055/s-0039-1688757] [Medline: 31117129]
63. Bickmore T, Mauer D, Crespo F, Brown T. Persuasion, Task Interruption and Health Regimen Adherence. In: Proceedings of the International Conference on Persuasive Technology. 2007 Presented at: PERSUASIVE'07; April 26-27, 2007; Palo Alto, CA, USA p. 1-11. [doi: 10.1007/978-3-540-77006-0_1]
64. MacDorman KF, Chattopadhyay D. Reducing consistency in human realism increases the uncanny valley effect; increasing category uncertainty does not. Cognition 2016 Jan;146:190-205 [FREE Full text] [doi: 10.1016/j.cognition.2015.09.019] [Medline: 26435049]
65. King AC, Bickmore TW, Campero MI, Pruitt LA, Yin JL. Employing virtual advisors in preventive care for underserved communities: results from the COMPASS study. J Health Commun 2013;18(12):1449-1464. [doi: 10.1080/10810730.2013.798374] [Medline: 23941610]
66. Abbott MB, Shaw P. Virtual nursing avatars: nurse roles and evolving concepts of care. Online J Issues Nurs 2016 Aug 15;21(3):7 [FREE Full text] [doi: 10.3912/OJIN.Vol21No03PPT39,05] [Medline: 27857172]
67. Langhorne P, Bernhardt J, Kwakkel G. Stroke rehabilitation. Lancet 2011 May 14;377(9778):1693-1702. [doi: 10.1016/S0140-6736(11)60325-5] [Medline: 21571152]
68. Nocera J, Hackney ME. The cognition-exercise interaction in Parkinson's disease: a perspective on current rehabilitative approaches with promise to impact the whole disease sequelae. J Geriatr Med Gerontol 2015 Sep 8;1:1-4. [doi: 10.24966/GGM-8662/100003]

Abbreviations

AI: artificial intelligence

COPD: chronic obstructive pulmonary disease
ECA: embodied conversational agent
e-coach: electronic coach
eCoaching: electronic coaching
eHealth: electronic health
e-learning: electronic learning
EU: European Union
IVR: interactive voice response
PD: Parkinson disease
RCT: randomized controlled trial
VC: virtual coaching
VNA: virtual nursing avatars

Edited by G Eysenbach; submitted 21.11.18; peer-reviewed by M Tielman, K Bul, M Alshehri; comments to author 31.03.19; revised version received 25.07.19; accepted 27.07.19; published 01.10.19.

Please cite as:
Tropea P, Schlieter H, Sterpi I, Judica E, Gand K, Caprino M, Gabilondo I, Gomez-Esteban JC, Busnatu S, Sinescu C, Kyriazakos S, Anwar S, Corbo M
Rehabilitation, the Great Absentee of Virtual Coaching in Medical Care: Scoping Review
J Med Internet Res 2019;21(10):e12805
URL: https://www.jmir.org/2019/10/e12805
doi:10.2196/12805
PMID:31573902

Review

Using Patient and Family Engagement Strategies to Improve Outcomes of Health Information Technology Initiatives: Scoping Review

Kevin Leung[1], MHI; Drew Lu-McLean[2]; Craig Kuziemsky[3], PhD; Richard G Booth[4], RN, PhD; Sarah Collins Rossetti[5], RN, PhD; Elizabeth Borycki[6], RN, PhD; Gillian Strudwick[1], RN, PhD

[1]Centre for Addiction and Mental Health, Toronto, ON, Canada
[2]McMaster University, Hamilton, ON, Canada
[3]University of Ottawa, Ottawa, ON, Canada
[4]Western University, London, ON, Canada
[5]Columbia University, New York, NY, United States
[6]University of Victoria, Victoria, BC, Canada

Corresponding Author:
Gillian Strudwick, RN, PhD
Centre for Addiction and Mental Health
1001 Queen St W
Toronto, ON, M6J1H4
Canada
Phone: 1 4165358501 ext 39333
Email: gillian.strudwick@camh.ca

Abstract

Background: Many health care organizations around the world have implemented health information technologies (ITs) to enhance health service efficiency, effectiveness, and safety. Studies have demonstrated that promising outcomes of health IT initiatives can be obtained when patients and family members participate and engage in the adoption, use, and evaluation of these technologies. Despite knowing this, there is a lack of health care organizations using patient and family engagement strategies to enhance the use and adoption of health ITs, specifically.

Objective: This study aimed to answer the following three research questions (RQs): (1) what current frameworks or theories have been used to guide patient and family engagement in health IT adoption, use, implementation, selection, and evaluation?, (2) what studies have been done on patient and family engagement strategies in health IT adoption, use, implementation, selection, and evaluation?, and (3) what patient and family engagement frameworks, studies, or resources identified in the literature can be applied to health IT adoption, use, implementation, selection, and evaluation?

Methods: This scoping review used a five-step framework developed by Arksey and O'Malley and adapted by Levac et al. These steps include the following: (1) identifying the RQ, (2) identifying relevant studies, (3) selecting studies, (4) charting relevant data, and (5) summarizing and reporting the result. Retrieved academic and grey literature records were evaluated using a literature review software based on inclusion and exclusion criteria by two independent reviewers. If consensus was not achieved, two reviewers would resolve conflicts by discussion. Research findings and strategies were extracted from the studies and summarized in data tables.

Results: A total of 35 academic articles and 23 gray literature documents met the inclusion criteria. In total, 20 of the 35 included studies have been published since 2017. Frameworks found include the patient engagement framework developed by Healthcare Information and Management Systems Society and the patient and family engagement framework proposed by Carman et al. Effective strategies include providing patients with clear expectations and responsibilities and providing reimbursement for time and travel. The gray literature sources outlined key considerations for planning and supporting engagement initiatives such as providing patients with professional development opportunities, and embedding patients in existing governance structures.

Conclusions: Several studies have reported their findings regarding successful strategies to engage patients and family members in health IT initiatives and the positive impact that can emerge when patients and family members are engaged in such initiatives in an effective manner. Currently, no framework has consolidated all of the key strategies and considerations that were found in

this review to guide health care organizations when engaging patients and family members in a health IT–specific project or initiative. Further research to evaluate and validate the existing strategies would be of value.

(J Med Internet Res 2019;21(10):e14683) doi:10.2196/14683

KEYWORDS

health information technology; informatics; patient engagement; medical informatics; health services

Introduction

Globally, health care organizations have implemented health information technologies (ITs) to improve health service efficiency, effectiveness, and patient safety [1-9]. Health IT in this context refers to several technologies implemented in clinical settings to improve an aspect of clinical care. These technologies may include but are not limited to (1) electronic health records (EHRs; medical records in electronic form), (2) patient portals (Web access to a variety of functions such as viewing lab results and booking an appointment), (3) care coordination portals (electronic portals that support the coordination of care activities), and (4) mobile health apps (apps to track, learn, or monitor an aspect of one's health) [10,11]. Some organizations have implemented patient portals that are tethered to hospital EHRs to provide patient access to their health information and provide patients with digital tools to take more control over their own health and care [12-14]. However, not all health care organizations have consistently taken full advantage of their health IT and systems [15]. Therefore, many of their potential benefits to patients, families, health professionals, and the health care system remain unseen [15].

Several studies have shown promising results from health IT when engaging end users such as patients and family members in the adoption, use, and evaluation of a specific health IT such as an EHR [16,17]. One article revealed a decrease in medication errors at the hospital when patients were engaged in the implementation of an electronic medication administration system [16]. By engaging patients in interviews to determine the best methods for identifying mental health patients through the barcode scanning process, an improvement in medication scanning rates and reduction in medication errors was realized [7]. In another study, patients engaged in user interface testing of a patient portal improved the usability and uptake of the technology by marginalized patient populations [18]. These examples show that active ways (eg, incorporating patients on design teams and obtaining patient feedback preimplementation of a health IT) of engaging patients and family members in health IT activities can be beneficial [18]. If health care organizations begin to effectively engage their patients and families in the development, adoption, use, and evaluation of health IT more regularly, there may be an increasing number of benefits realized.

Engaging patients and family members in decision making related to health services is not a new concept [19]. In fact, patients have engaged with health care organizations in many ways for several years now, such as advising health care institutions about health service delivery and patient-oriented health research [20]. In the United States, the Patient-Centered Outcomes Research Institute has advocated and funded the engagement of patients in many research initiatives [21]. However, patient engagement in health IT development, adoption, use, and evaluation is not as mature as it is in patient participatory medicine for health service delivery. The term, *patient engagement*, in this context broadly encompasses processes that involve informing and/or partnering with patients [22], to inform health service planning and delivery of the health IT. Engaging in a patient's own health care is outside of the scope of this discussion.

The objective of this scoping review was to explore studies that aim to improve outcomes of health IT initiatives through patient and family engagement and to explore practical strategies that health care organizations can leverage to meaningfully engage patients and families across the continuum of health IT initiatives. The review specifically focuses on how patients and family members can be involved in the health service planning and delivery context and not how they are more engaged in their own care.

Methods

This scoping review adopted the 5-step framework defined by Arksey and O'Malley [23] and updated by Levac et al [24]. These steps include the following: (1) identifying the research question (RQ), (2) identifying relevant studies, (3) selecting studies for data collection, (4) charting relevant data, and (5) summarizing and reporting the result [23,24]. Ethical approval was not required for this study. The following sections outline the steps followed to conduct the review.

Step 1: Identifying the Research Questions

The literature review was conducted to answer the following three RQs:

- RQ1: What current frameworks, models, or theories have been used to guide patient and family engagement in health IT adoption, use, implementation, selection, and evaluation?
- RQ2: What studies have been done on patient and family engagement strategies in health IT adoption, use, implementation, selection, and evaluation? And what are their results?
- RQ3: What patient and family engagement frameworks (not specific to health IT), studies, and/or resources can be applied to health IT adoption, use, implementation, selection, and evaluation?

The use of the 3 RQs was done to clearly articulate the scope of the review and explore literature that can be applied broadly to patient and family engagement in health IT projects or initiatives. The following assumptions were made to clarify the terminology used in the study when developing the RQs. The

broader term *health IT* was used to represent various forms of IT used in health contexts such as (but not limited to) *EHRs* (patient records accessible via a computer) [25], *patient portals* (secure websites that allow patients to access their health record and other functions such as booking a medical appointment) [26], *care coordination portals*, and *mobile health apps* (use of mobile devices to provide health care services) [27]. RQs and terminology were refined through consultation with the members of the research team and a patient and family advisory committee. The patient and family advisory committee was based at a Canadian hospital located in Toronto, Ontario.

Step 2: Identifying Relevant Studies

With the guidance of a research librarian with experience conducting scoping reviews, a search strategy was developed using the following databases: MEDLINE, PsycINFO, Culmulative Index for Nursing and Allied Health Literature, Theses Canada, and the Education Resources Information Center. These databases are all commonly used in health science–related literature reviews with a focus on health services and were made available through the organizations in which the authors are employed. A primary search strategy was developed for the MEDLINE database (see Table 1) and adapted to be used for the other electronic databases.

The search strategy used a combination of relevant Medical Subject Headings terms identified by the research librarian and keywords with Boolean terms (search combinations are shown in Table 1 via the use of the various search syntax shown above). The database search was supplemented with a search for unpublished gray literature related to patient engagement frameworks and toolkits. The search tool by the Canadian Agency for Drugs and Technologies in Health, Grey Matters, guided the gray literature search. A search using Google's search engine was also completed. Additional methods for identifying relevant resources included the following: (1) reference searching from key article reference lists, (2) input from health IT experts, and (3) the patient and family advisor committee.

The gray literature search used keywords to refine the search that included the following: ehealth, electronic medical records, electronic health records, patient portals, patient/families/caregiver engagement, ehealth adoption, use and evaluation, frameworks/ strategies/ resources/ tools/ toolkits, and health IT. Searches were completed between June and December 2018, and the screening process took part in early 2019.

Inclusion/Exclusion Criteria

Studies and frameworks in any clinical or health care setting were included. The studies and documents assessed were not limited to their date of publication or country of origin. Studies not published in English were excluded from this review. Systematic reviews were not eligible, but their reference lists were screened to find supplementary relevant studies. Studies that did not explicitly address patient or family engagement were excluded from the review. The focus of the review was on studies that involved, engaged, or empowered patients, families in the decision-making process, and across all stages of a health IT initiative for health service planning and delivery and not studies that focused on technologies that can be used to engage patients, families, and/or caregivers in their care. Studies that focused on patient engagement as an outcome of integrating and enabling technologies in the care process were, therefore, excluded.

Table 1. Shows the search strategy developed for the MEDLINE database.

No	Searches	Results
1	health records, personal/or patient portals/	1531
2	exp Medical Records Systems, Computerized/ or exp Electronic Health Records/ or exp Hospital Information Systems/ or exp Information Systems/	235,853
3	exp Patient Participation/	23,426
4	((patient or family or caregiver) adj2 (engag* or involv* or empower* or activat* or participat*)).mp.	46,576
5	framework.ab,ti.	217,304
6	exp Telemedicine/	24,262
7	exp Patient Portals/	179
8	"patient and public involvement".kw.	145
9	exp Medical Informatics/	419,180
10	theory.ab,ti.	303,436
11	model.ab,ti.	1,870,340
12	Search #5 or Search #10 or Search #11 [framework set]	2,259,697
13	3 or 4 or 8 [patient engagement set]	46,658
14	Biomedical Technology/	5923
15	Search #1 or Search #2 or Search #6 or Search #7 or Search #9 or Search #14 [eHealth set]	472,799
16	Search #12 and Search #13 and Search #15	558

Step 3: Study Selection

Numerous records were retrieved from the academic database search and gray literature search. Literature screening was completed using the Covidence systematic review software [28]. The software removed duplicates from the database search to aid with screening. Once duplicates were removed, 2 members of the research team independently screened the article titles and abstracts to determine if the full-text article should be retrieved and assessed. After the title and abstract screening, the inter-rater reliability score was calculated. Conflicts/discrepancies between screeners were resolved by discussion between the 2 screeners, and if consensus was not reached, a third member of the research team was consulted.

Step 4: Charting the Data

From each included study or gray literature document, pertinent information was extracted and summarized to address the RQs. This information included the following: descriptions of the study (study name, authorship, country of publication, journal published, study design, and population study), study methods (engagement strategies employed by researchers), and all study results related to the RQs (lessons learned and recommended engagement strategies).

Step 5: Collating, Summarizing, and Reporting Results

Results were collated, summarized, and reported based on the RQ or RQs that the article addressed. Descriptive statistics were used to summarize characteristics of the articles found through the database and gray literature search. A content analysis was performed on the identified studies to identify and record overarching themes that emerged. The approach to categorizing themes was an iterative, inductive process involving 2 members of the research team. The 2 members of the research team met after the data extraction process to create and refine the overarching themes identified. The recommended patient engagement strategies and considerations discussed in the included articles were also recorded.

Results

Characteristics of the Identified Studies

A total of 1395 academic literature records, or gray literature documents, were retrieved from the academic and gray searches and expert consultation. During the abstract and title screening process, the inter-rater reliability between the 2 independent raters (KL/DM) was above 75% (Cohen kappa=.44). After the full-text records were assessed, 35 (n=35) academic articles and 23 (n=23) gray literature documents met the inclusion criteria (see Figure 1). The publication date of the academic articles and gray literature ranged between 2005 and 2018. The studies were conducted in the United States (n=14), and the remaining publications originated in other countries located in North America, Europe, and Asia. From the gray literature, 13 (n=13) documents were published in Canada, 4 (n=4) in Australia, 3 (n=3) in the United Kingdom, and 3 (n=3) in the United States. Table 2 provides an overview of the design of the academic and gray literature included in this review. Multimedia Appendix 1 provides the full data extraction table for all included records.

Figure 1. Preferred Reporting Items for Systematic Reviews and Meta-Analyses diagram indicating the flow of records reviewed during the literature search.

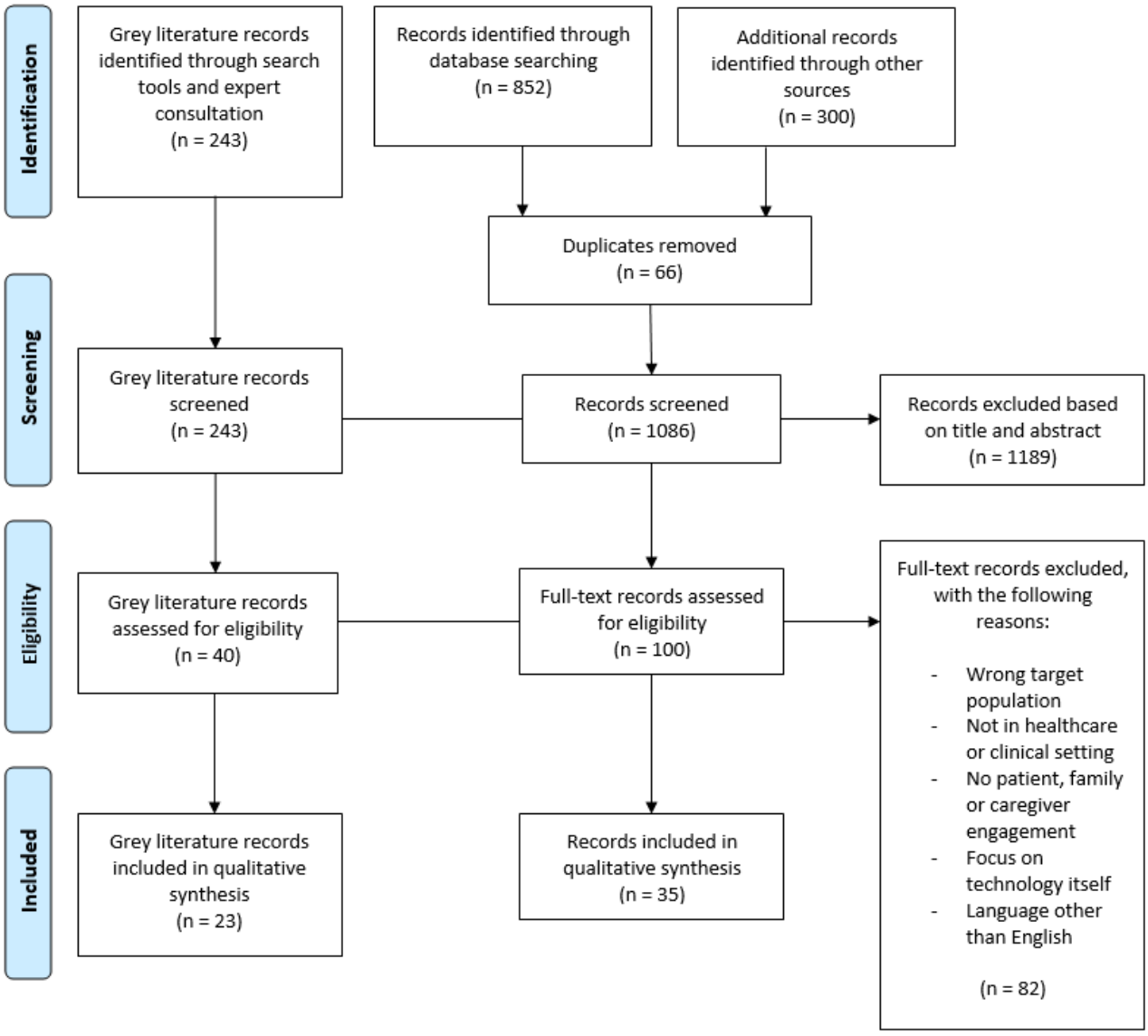


Table 2. Design of the identified academic and gray literature (N=58).

Design	Statistics, n (%)
Quantitative	1 (2)
Qualitative	21 (36)
Mixed methods	1 (2)
Other academic papers (eg, case reports and editorials)	12 (20)
Reports/websites/other forms of gray literature	23 (40)

RQ1: Current Health Information Technology–Related Frameworks, Models, or Theories

A total of 2 articles outlined frameworks or models that have been used to effectively guide patient and family engagement in health IT adoption, use, implementation, selection, and evaluation [29,30]. In the first article that addresses RQ1, the Healthcare Information and Management Systems Society patient engagement framework was discussed. This framework was created to inform health care organizations on how to leverage health IT systems to implement patient engagement strategies [29]. The framework outlines 5 approaches to engagement that align with the International Association for Public Participation's spectrum for community engagement and suggest tools such as patient-accessible records and patient-specific education to inform and partner with patients. Although these tools are not at the health service delivery level, the stages of engagement discussed in the framework are applicable to health IT health service planning and delivery.

These stages of engagement are as follows: (1) inform, (2) consult, (3) involve, (4) collaborate, and (5) empower [22].

The second article that addressed RQ1 is a study by Walker et al that outlined an evaluative model for health IT implementation [30]. Walker et al engaged patient groups through a series of focus groups and an online forum to understand the evaluation needs of patients related to an inpatient portal implementation. Through a nominal group approach, patients specified the following 2 categories of patient-specific outcomes that should be incorporated in an evaluation of an inpatient portal: (1) outcomes related to the explicit use of the inpatient portal and (2) outcomes related to the tacit knowledge gained by having access to an inpatient portal [30]. Explicit use of the inpatient portal may result in a change in patient satisfaction, which can be a measure of patient outcomes as a result of portal implementation and use [30]. Tacit knowledge that is gained from using the inpatient portal could result in change in patient engagement in health IT, which can be a measurable outcome [30]. The model demonstrates that soliciting patient feedback on explicit portal use and changes in tacit knowledge is needed for a multifaceted evaluation of an inpatient portal implementation.

RQ2: Patient and Family Engagement Strategies in Health Information Technology Initiatives

A total of 19 (n=19) studies were identified that utilized patient and family engagement strategies in health IT adoption, use, implementation, selection, and evaluation. A finding of 2 studies related to the use of engaging patients through multiple digital modalities. A study by Athilingam et al engaged patients on the design, potential features, and the ease of use of a mobile health app to improve self-care in heart failure [31]. Interviews were conducted with 10 patients and 4 cardiologists before the initial prototype was developed. The authors concluded that existing mobile health apps have not been widely adopted and suggested the mixed use of internet, email, and text messages for promoting better communication and long-term engagement with digital health apps [31]. Another study by Lafata et al reiterated the need for different digital modalities to effectively engage patients from vulnerable patient populations [32].

Multidisciplinary, Team-Based Approach

A total of 4 (n=4) studies suggested that a multidisciplinary, team-based approach would be an effective engagement strategy for engaging patients and families in the use of health IT [33-36]. A study by Ackerman et al examined the patient engagement strategies during the implementation of a patient portal at 5 community health centers [33]. Portal champions reminded clinicians and staff to encourage patients to sign up for the portal. Volunteers, front desk clerks, and medical assistants provided enrollment assistance to patients and used clinic computers to demonstrate to patients how to use portal services. The study by Ackerman et al found that the uptake and use of the portal improved when patients were engaged by trusted staff members or clinicians. Results from the study by Raval et al concluded that the engagement of a pediatric surgeon and physician assistant was crucial to the success of recruiting and engaging patients in the development of a mobile app for colorectal disease management [34]. As the clinical team was available during the planning stages of the research project, it helped ease the process of engaging patients in the study. As a result, pediatric patients and their family members who were already visiting the hospital were successfully recruited and provided useful feedback. Similar recommendations were made by Krist et al and Shapiro-Mathews et al [35,36]. In a study by Krist et al, patient engagement strategies in primary care practices were evaluated [35]. Primary care offices had registration staff pass out information to patients, had nurses discuss how to sign up for a patient portal, and other clinicians communicate the value of signing up for the patient portal. Krist et al concluded that a team-based approach to engaging patients positively influenced the uptake of the patient portal compared with the clinician-dependent approach to engaging patients. The article by Shapiro-Mathews et al outlined an institutional strategy for mobile health technologies that requires an interdisciplinary approach [36]. A clinical nurse specialist and other health care professionals can facilitate a team-based approach to engage patients, provide patient education, and inform the design of mobile apps that meet the needs of patients.

Training/Education of Patients

A total of 8 (n=8) articles highlighted the importance of training/education of patients in the success of health IT adoption, use, implementation, selection, and evaluation. The results from a study by Anshari et al showed that the availability of a Web-based health educator is important to improving the health literacy of patients and empowering patients to control their own health and health information [37]. Greysen et al conducted a randomized controlled trial where patients in the intervention arm received tailored, structured education regarding the use of a patient portal at the bedside [38]. Study results suggested that bedside portal training produced a trend of increased ability to log in and navigate the portal, satisfaction with portal use, and frequency of portal use after discharge. Van den Bulck et al explored patient needs, expectations, and attitudes toward a patient portal by administering an online survey to recruited patients [39]. Results showed that digital health literacy is a key factor to portal adoption and providing training to patients could provide exposure to using the portal and create appropriate expectations of what the portal is capable of. Another study by Wildenbos et al explored experiences and perspectives of older adult patients using a patient portal 1-year postimplementation of the technology [40]. The results from the study by Wildenbos et al concluded that health literacy level of patients was a strong factor that influenced the patient's overall interest and perceived ability to use the patient portal. Although the previously discussed training/education methods may be most directed at engaging patients in their own care, similar methods could be used to orient patients before engaging in a health IT initiative aimed at supporting adoption, use, implementation, selection, and evaluation.

Training of Health Care Providers

In addition to training patients on the use of health IT, the study by Raval et al also recommended that health care providers may need to be trained to address potentially low levels of self-efficacy of employing effective patient engagement methods [34]. Wildenbos et al also emphasized the need for health care

providers to be trained regarding how to effectively engage patients [41]. Likewise, a study by Metting et al explored patient needs and opinions through focus groups to facilitate the development of patient Web portals [42]. A training activity that was recommended included training health care providers to effectively communicate with patients with regards to engagement specific activities.

RQ3: General Patient and Family Engagement Frameworks, Studies, and Resources

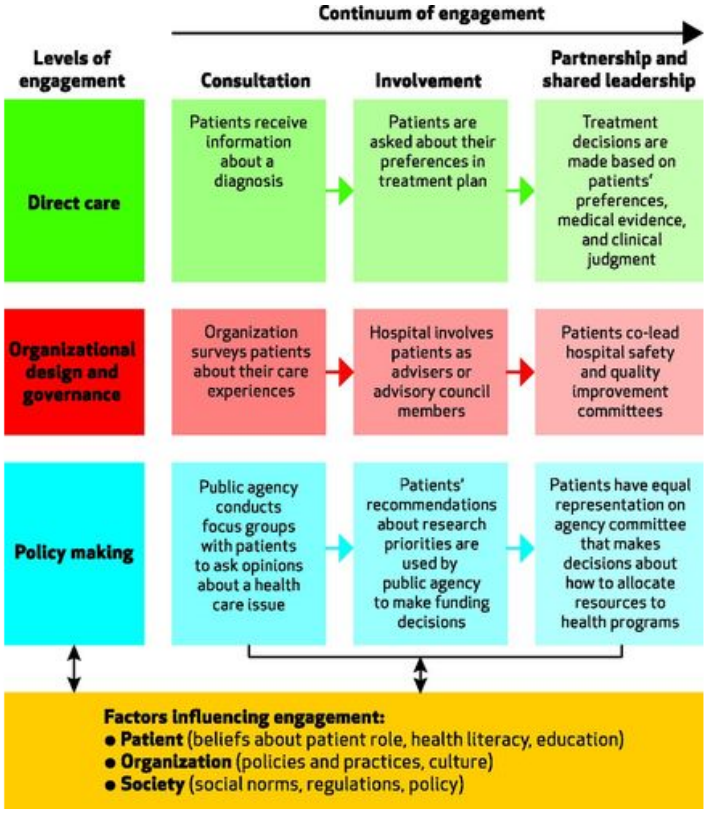
A total of 13 (n=13) academic sources and 23 (n=23) gray literature sources were found that employed patient and family engagement strategies for research and clinically relevant projects that were not health IT–specific, but the principles or findings embedded within them could potentially be applied to this context. In turn, relevant principles and findings may be included in a future health IT–specific patient and family engagement framework or set of recommendations for health care organizations to utilize. Of the 13 academic articles, 6 (n=6) explicitly outlined frameworks that can be used to involve patients and family members within nonhealth IT contexts such as research, health service delivery, and quality improvement.

A framework proposed by Carman et al was widely used to inform effective patient and family engagement strategies in health IT adoption, implementation, use, selection, and evaluation [43] (see Figure 2). The framework outlines 3 categories of engagement activities across a continuum: (1) consultation, (2) partnership, and (3) shared leadership. These 3 categories can be applied in different levels of the health care system, which was segmented by the authors as follows: individual care, organization governance, and government policy. The framework highlighted individual factors that can potentially impact a patient's willingness and ability to engage with the health care system such as health literacy and education level. At the institutional level, health care organizations and staff can also encourage patient engagement through demonstrating that patient participation and leadership is imperative to the achievement of organizational goals [43].

Policies and practices that can influence patient engagement can create expectations that patients can and will serve as advisors and decision makers on committees and patient-centered councils. Bridgepoint Hospital in Toronto, Canada, implemented practices that brought patients, families, and hospital staff together to redesign health care services at the hospital and improve the overall patient experience [44]. Patients and family members at the hospital were recruited as advisors to review the hospital quality improvement processes.

At the government level, policy makers can create mechanisms so that patients can provide input in public policy, such as public deliberation sessions, town hall meetings, public hearings, or regulatory comment processes [43]. Nonprofit organizations and government agencies can also aid in creating funding mechanisms requiring patient participation in the decision-making process. These efforts to encourage patients and family member participation in quality improvement processes broadly can be applied specifically to the health IT context, creating mechanisms that encourage patients and families to provide their input in health IT–related projects and policies.

Figure 2. Multidimensional framework for patient and family engagement in health and health care (permission from author and publisher obtained to use image, and copyright obtained).



A number of the identified studies that addressed RQ3 utilized different strategies to engage patients in a clinical research network or to engage the youth in projects led by academic researchers (see Table 3).

Guidance documents and reports identified in the gray literature highlighted key attributes to patient engagement broadly in the health care and research context. Gray literature sources highlighted key considerations to guide health care organizations in preparing for engagement activities (see Table 4).

Table 3. A list of common strategies used in identified academic studies to engage patients and family members.

Recommended strategies	Identified studies
Provide stakeholders with clear expectations, roles, and responsibilities (eg, number of hours and anticipated deliverables)	Perfetto et al (2018) [19], Arkind et al (2015) [45], Boyer et al (2018) [46], Hawke et al (2018) [47], Hamilton et al (2018) [48], Shelef et al (2018) [49], Warren et al (2018) [50]
Develop policy or practice that provides incentive or compensation to stakeholders for their time and efforts (food, travel, reimbursement for time, and provision of training opportunities)	Perfetto et al (2018) [19], Arkind et al (2015) [45], Hawke et al (2018) [47], Hamilton et al (2018) [48], Shelef et al (2018) [49], Warren et al (2018) [50]
Engage stakeholders early in the planning and development stage of the project	Perfetto et al (2018) [19], Arkind et al (2015) [45], Boyer et al (2018) [46], Chung et al (2018) [51], Faulkneret al (2018) [52]
Be transparent about patient contributions being used and making an impact on the project	Perfetto et al (2018) [19], Hawke et al (2018) [47], Chung et al (2018) [51], Coathup et al (2016) [53]
Prioritize effective communication with regular updates and provide explanation of research/medical terminology and show that patients are being valued as partners	Boyer et al (2018) [46], Chung et al (2018) [51], Faulkneret al (2018) [52], Coathup et al (2016) [53]
Allow stakeholders to meet in a convenient time (eg, weekday evenings) and location (eg, remote access)	Perfetto et al (2018) [19], Arkind et al (2015) [45], Hawke et al (2018) [47], Hamilton et al (2018) [48]
Engage a group of more than 2 patients so that they can support one another and have shared discussion	Perfetto et al (2018) [19], Hamilton et al (2018) [48], Faulkneret al (2018) [52]
Leverage clinical providers as trusted agents	Boyer et al (2018) [46], Chung et al (2018) [51], Suarez-Balcazar (2005) [54]
Provide adequate preparation (orientation, training, and resources) for both the engaged stakeholders and the team members engaging stakeholders	Boyer et al (2018) [46], Suarez-Balcazar (2005) [54]
Use established networks of stakeholder groups	Perfetto et al (2018) [19], Boyer et al (2018) [46]
Identify concepts that may be confusing and set aside time to explain them in jargon-free terms	Hawke et al (2018) [47], Hamilton et al (2018) [48]

Table 4. Key considerations from existing engagement frameworks and toolkits when preparing for engagement.

Considerations when preparing for engagement	Identified studies
Clarify objectives and impact of engagement activities	[55-57]
Clarify why stakeholders can get involved	[21,58-61]
Discuss preferences of ongoing communication to ensure that all stakeholders involved are informed throughout	[62,63]
Determine the appropriate level of engagement that meets patient and organizational goals	[64]
Clarify and document roles, responsibilities, and scope of engaged population that are agreed upon between project team and engaged population	[58,59,63,65]
Define time commitment and expectations from the engagement team	[55,56,58,60,62]
Consider potential barriers for engagement	[62]
Plan to protect patient privacy	[56]
Involve patients in the planning process	[21,56,57,59,64,66]
Allocate financial resources to reimburse participants for expenses incurred while being involved in project (eg, time, travel, training, translation, and childcare)	[56,58-60,63,67,68]

Additional considerations and characteristics were highlighted to support patients, family members, and staff in health care organizations during engagement activities (see Table 5). In addition to preparing and supporting engagement, 9 (n=9) gray literature sources highlighted the importance of conducting an evaluation to ensure that initiatives to engage patients and family members are meaningful and provide value for all stakeholders [21,57,59,61-65,67]. Studies encouraged the use of quantitative

measures and measuring against specified objectives to demonstrate the value of engagement [57,67]. Different evaluative methods were described such as providing evaluation forms, conducting surveys, and engaging in discussions with patients to provide the opportunity for giving constructive feedback. The use of surveys and feedback forms allows for organizations to solicit feedback anonymously, which may be preferred by some patients and family members [62,63,69].

Table 5. Key considerations to support patients and family members during engagement activities outlined by gray literature sources.

Considerations when supporting engagement activities	Identified studies
Meet patients/families in their own environments and communities	[59,60]
Schedule meetings at a variety of different times (not just during working hours)	[59]
Allow for family presence when engaging patients	[66]
Engage 2 or more patients at a time	[56]
Provide training and support for stakeholders to effectively communicate and partner with each other	[58,60,62]
Leverage Web-based tools when providing training for patient engagement	[68]
Ensure researchers/project leads have the necessary skills to involve patients	[58,59]
Provide patients with background information/readings before preparing for meetings	[56,63]
Provide patients with opportunities for professional development (eg, authoring manuscripts and presenting at conferences)	[21,60,62,67]
Embed patient partners in existing governance structures (eg, boards, steering committees, advisory councils, and patient groups)	[55,64,67,70]
Communicate in jargon-free language	[57]
Track and update clear timelines of each milestone	[68]
Leverage health care professionals and their ability to encourage patients to engage	[71]
Clearly articulate the type of patient information being shared and why	[56]
Be transparent about the constraints and why input may not be always used	[58,59,61,63,68,72]
Outline outcomes that are important to patients and can improve their quality of life	[55]
Frequently check in with patients for any questions and to keep them informed (eg, through progress reports or newsletters)	[62]
Provide opportunity for stakeholders to use multiple channels of communication (eg, written communication, emails, phone calls, and social media)	[59,63,70]
Provide thank you letters, along with feedback and suggestions for future involvement	[68]

Discussion

Primary Findings

Patient and family engagement has become a topic of increasing interest in the research community in the last several years. In fact, the majority (57%, 33/58) of the included studies in this review have been published since 2017. Identified studies have highlighted the importance of involving and partnering with patients in health IT design and development [51,73]. Several studies have used a variety of strategies and outlined considerations when engaging patients and family members in health IT–related projects. These studies outlined several considerations for engaging patients and families, mainly when preparing for engagement or in the early stages of the engagement process (see Table 4). Strategies that were consistent among the identified studies include providing incentives and reimbursement for patients, clarification of patient roles and responsibilities, and demonstrating the value of engagement to patients. However, there is a limited amount of literature that have established a framework, model, or theory, or set of recommendations that can be used to guide patient and family engagement in health IT initiatives. As a result, there is no standardized methodology or resource for engaging patients and family members effectively for health IT initiatives. The development of a practical framework (borrowing concepts, recommendations, and relevant principles from the nonhealth IT literature) to guide health care organizations and health care providers within this unique context is needed. One main limitation for the development of a standardized framework is that it may not capture the unique needs of all diverse patient groups. Therefore, additional resources may be required to supplement and guide effective patient and family engagement for health IT initiatives; however, starting with a resource that health care organizations can use to engage in this important work, is needed.

Patient engagement is a broad term that can be defined by varying levels of involvement from the patient. The overall goal of patient engagement may not always be about moving to a higher level of engagement and patient empowerment. The study by Hamilton et al outlined a key recommendation of

patients valuing the freedom to become gradually more engaged at their own pace. Therefore, desired objectives and the impact of engagement need to be discussed and agreed upon between the organization and the engaged stakeholders. Many health care organizations recognize the varying degrees of engagement. A total of 9 (n=9) citations found in this review referred and/or adapted the participation spectrum outlined by the International Association for Public Participation that identifies different levels of stakeholder and community participation [22]. These findings are relevant not only in the broader health care or health research settings but can be applied to health IT contexts.

Comparison With Past Work

This scoping review adds to existing literature reviews that have articulated successful engagement characteristics and approaches within the health research context [74]. The impact of engagement can be profound. The review by Manafo et al documented outcomes when engaging patients in health research that include the following: (1) patients feeling empowered, (2) improved trust between researcher and patients, and (3) decreased attrition of study participants. Increased engagement has also positively impacted health IT projects in different ways such as increased usage of health IT, satisfaction with health IT use [35,37,38], and the obtainment of quality/safety-related outcomes [17]. A review on patient involvement in health research by Shippee et al concluded that available literature on patient involvement focused on 1 research phase and particularly earlier stages of research. Similarly, academic studies have involved patients in specific stages of health IT projects, such as the usability testing or the implementation of an inpatient portal and not necessarily throughout all stages of the health IT life cycle. An article by Petersen articulated that despite patients being actively involved in conducting research, patients are often not involved with setting the research agenda, evaluating results, or discussing next steps [20]. Similarly, there is limited evidence reporting or evaluating engagement strategies across multiple phases of a health IT project. As a result, there is a need for further research on sustained engagement throughout different stages of health IT projects and how strategies may differ depending on whether the organization is in a development or implementation phase of a health IT project. This is where identifying findings from the nonhealth IT literature may add value.

Studies have highlighted the important role of health providers in the engagement and activation of patients [57,65]. A study by Graffigna concluded that health providers have a crucial role in influencing the engagement and activation of type 2 diabetic patients in using health IT tools to manage their health condition [75]. Furthermore, several academic and gray literature sources have highlighted the need to incorporate patients and families in underserved populations in health IT initiatives. Involvement and feedback that reflects the diversity of the community can allow organizations to gain a better understanding of the diversity in patient needs [59,68,76,77]. Studies suggested that there were economic and ethnic disparities associated with the use and adoption of health IT [78,79]. Future studies could explore methods in which current engagement strategies can be adapted to effectively engage patients and families from underserved populations. There is also opportunity for health care organizations that primarily work with underserved populations to document and disseminate their strategies in effectively engaging these populations.

Gray literature sources have outlined the importance of evaluating engagement activities to quantify the value of engagement and provide constructive feedback on how engagement initiatives can be improved [63,67]. Evaluation methods that have been recommended include having feedback discussions with patients, providing evaluation forms, and surveys. Other sources have also considered standardizing evaluation by building evaluation components into project plans and leveraging existing standardized tools to evaluate the process and impact of engagement [56,59]. A study by Abelson et al developed a patient engagement evaluation tool that can be used by health care organizations broadly. The evaluation tool consists of 3 unique questionnaires used to evaluate the following 4 evaluation principles: (1) integrity of design and process, (2) influence and impact, (3) participatory culture, (4) collaboration, and (5) common purpose [80].

Despite the number of tools that have been developed to evaluate public and patient engagement, there is a lack of consistency on how engagement strategies are evaluated, and few studies have quantitatively evaluated measures of engagement. A literature review by Esmail et al outlined that there were only 2 studies that evaluated patient engagement within the health research context using quantitative measures [81]. A review by Dukhanin et al identified other existing evaluation tools but concluded that the methodology of each tool varies significantly [82]. Therefore, implications for research include the validation of existing evaluation tools and the combined use of qualitative and quantitative tools to assess engagement. Furthermore, many existing evaluation tools rely heavily on process metrics and measure perceived benefits of engagement [82]. There is a need for health care organizations to increase adoption and use of outcome metrics, such as changes in patient knowledge and attitudes toward engagement, to evaluate patient, public, and community engagement. Meaningful evaluation efforts require capacity and commitment from health care organizations [80]. An implication for health care organizations is to internally build the capacity and culture that supports evaluation efforts to improve the engagement process for patients, families, and organizations.

Limitations

This review has a few limitations that should be considered when reviewing its findings. The exploration of engagement strategies broadly has led to several considerations and recommendations for how to engage patients and families in health IT initiatives. As identified in the literature, there is a spectrum that exists within the context of patient engagement. Patients can be actively or passively involved in health IT initiatives, and this distinction should be further explored throughout the stages of engagement within the health IT context. A more focused approach that applies to specific health care settings may be completed to solicit results that are appropriate for specific contexts.

Regarding the review methodology, there were challenges with the search strategy regarding the use of broad terms such as

theory, model, and frameworks. The terms theory, model, and frameworks were taken from common terminology and approaches used in implementation science [83]. This was done to capture as many different types of engagement tools that could guide future engagement strategies; but as a result, there were a significant number of studies that did not meet the appropriate inclusion criteria for this review but were identified in the initial search. The review included studies published in English only. There is a possibility that relevant studies and resources from health care organizations in non-English speaking countries that have studied or developed engagement strategies in health IT settings were excluded.

Conclusions

Several studies and gray literature documents identified in this review have reported their findings on successful strategies to engage patients and family members in health care and the positive impact that can emerge when patients and family members are engaged. Several studies have employed a variety of engagement strategies to engage diverse patient populations in health IT projects. Currently, no framework, set of recommendations, or resource document has consolidated all of the lessons learned and considerations to guide health care organizations when engaging patients and family members in a health IT–specific project. There is much to learn and incorporate from the nonhealth IT–specific work that has already been done. With the increasing number of studies reporting their findings of engaging patients in health IT–related initiatives, continuing efforts to evaluate and validate these engagement strategies is needed.

Acknowledgments

The authors would like to acknowledge the patient and family advisory committee for their contributions in conducting this scoping review. The authors would also like to thank the research librarian who provided guidance for the search strategy utilized. Finally, the authors would like to thank the Canadian Institutes of Health Research for funding this study.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Academic literature data extraction table.
[PDF File (Adobe PDF File), 205 KB - jmir_v21i10e14683_app1.pdf]

References

1. Better Information for Improved Health: A Vision for Health System Use of Data in Canada. Ottawa, ON: Canadian Institute for Health Information; 2013.
2. Chang F, Gupta N. Progress in electronic medical record adoption in Canada. Can Fam Physician 2015 Dec;61(12):1076-1084 [FREE Full text] [Medline: 27035020]
3. Powell KR, Myers CR. Electronic patient portals: patient and provider perceptions. Online J Nurs Informatics 2018;22(1):1 [FREE Full text]
4. Schoen C, Osborn R, Squires D, Doty M, Rasmussen P, Pierson R, et al. A survey of primary care doctors in ten countries shows progress in use of health information technology, less in other areas. Health Aff (Millwood) 2012 Dec;31(12):2805-2816. [doi: 10.1377/hlthaff.2012.0884] [Medline: 23154997]
5. Institute for e-Health Policy. 2014. A Glimpse at EHR Implementation Around the World: The Lessons the US Can Learn URL: https://www.e-healthpolicy.org/sites/e-healthpolicy.org/files/A_Glimpse_at_EHR_Implementation_Around_the_World1_ChrisStone.pdf [accessed 2019-05-08]
6. Gray BH, Bowden T, Johansen I, Koch S. Electronic health records: an international perspective on 'meaningful use'. Issue Brief (Commonw Fund) 2011 Nov;28:1-18. [Medline: 22164356]
7. Sulkers H, Tajirian T, Paterson J, Mucuceanu D, MacArthur T, Strauss J, et al. Improving inpatient mental health medication safety through the process of obtaining HIMSS Stage 7: a case report. JAMIA Open 2019;2(1):35-39. [doi: 10.1093/jamiaopen/ooy044]
8. Kipping S, Stuckey MI, Hernandez A, Nguyen T, Riahi S. A web-based patient portal for mental health care: benefits evaluation. J Med Internet Res 2016 Nov 16;18(11):e294 [FREE Full text] [doi: 10.2196/jmir.6483] [Medline: 27852556]
9. FitzHenry F, Wells N, Doran J, Hughart K, Levy M, Doulis J. Applying bar code medication administration to make a difference in adverse drug events with potential for harm: lessons learned. Comput Inform Nurs 2013 Oct;31(10):457-462. [doi: 10.1097/CIN.0000000000000020] [Medline: 24135882]
10. OpenNotes – Patients and Clinicians on the Same Page. 2018. URL: https://www.opennotes.org/ [accessed 2018-02-08]
11. Patel V, Johnson C. HealthIT. 2018. Individuals' Use of Online Medical Records and Technology for Health Needs URL: https://www.healthit.gov/sites/default/files/page/2018-03/HINTS-2017-Consumer-Data-Brief-3.21.18.pdf [accessed 2019-09-04]

12. HealthIT. 2017. What is a Patient Portal? URL: https://www.healthit.gov/providers-professionals/faqs/what-patient-portal [accessed 2019-09-05]
13. Irizarry T, DeVito DA, Curran CR. Patient portals and patient engagement: a state of the science review. J Med Internet Res 2015 Jun 23;17(6):e148 [FREE Full text] [doi: 10.2196/jmir.4255] [Medline: 26104044]
14. Irizarry T, Shoemake J, Nilsen ML, Czaja S, Beach S, DeVito Dabbs DA. Patient portals as a tool for health care engagement: a mixed-method study of older adults with varying levels of health literacy and prior patient portal use. J Med Internet Res 2017 Mar 30;19(3):e99 [FREE Full text] [doi: 10.2196/jmir.7099] [Medline: 28360022]
15. Government of Canada Publications. 2010. Electronic Health Records in Canada: An Overview of Federal and Provincial Audit Reports URL: http://publications.gc.ca/site/eng/367717/publication.html [accessed 2019-09-05]
16. Paoletti RD, Suess TM, Lesko MG, Feroli AA, Kennel JA, Mahler JM, et al. Using bar-code technology and medication observation methodology for safer medication administration. Am J Health Syst Pharm 2007 Mar 1;64(5):536-543. [doi: 10.2146/ajhp060140] [Medline: 17322168]
17. Strudwick G, Clark C, McBride B, Sakal M, Kalia K. Thank you for asking: exploring patient perceptions of barcode medication administration identification practices in inpatient mental health settings. Int J Med Inform 2017 Sep;105:31-37. [doi: 10.1016/j.ijmedinf.2017.05.019] [Medline: 28750909]
18. Eyasu T, Leung K, Strudwick G. Guiding improvements in user experience: results of a mental health patient portal user interface assessment. Stud Health Technol Inform 2019;257:110-114. [Medline: 30741182]
19. Perfetto EM, Harris J, Mullins CD, dos Reis S. Emerging good practices for transforming value assessment: patients' voices, patients' values. Value Health 2018 Apr;21(4):386-393 [FREE Full text] [doi: 10.1016/j.jval.2017.11.013] [Medline: 29680093]
20. Petersen C. Patient informaticians: turning patient voice into patient action. JAMIA Open 2018;1(2):130-135. [doi: 10.1093/jamiaopen/ooy014]
21. Patient-Centered Outcomes Research Institute (PCORI). 2019. Improving outcomes important to patients URL: https://www.pcori.org/ [accessed 2019-07-13]
22. IAP2 Canada. 2019. IAP2 Spectrum URL: https://iap2canada.ca/Resources/Documents/0702-Foundations-Spectrum-MW-rev2%20(1).pdf [accessed 2019-07-13]
23. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodol 2005 Feb;8(1):19-32 [FREE Full text] [doi: 10.1080/1364557032000119616]
24. Levac D, Colquhoun H, O'Brien KK. Scoping studies: advancing the methodology. Implement Sci 2010 Sep 20;5:69 [FREE Full text] [doi: 10.1186/1748-5908-5-69] [Medline: 20854677]
25. Häyrinen K, Saranto K, Nykänen P. Definition, structure, content, use and impacts of electronic health records: a review of the research literature. Int J Med Inform 2008 May;77(5):291-304. [doi: 10.1016/j.ijmedinf.2007.09.001] [Medline: 17951106]
26. Canada Health Infoway: Digital Health in Canada. 2019. Electronic Health Records URL: https://www.infoway-inforoute.ca/en/solutions/digital-health-foundation/electronic-health-records [accessed 2019-07-13]
27. Free C, Phillips G, Galli L, Watson L, Felix L, Edwards P, et al. The effectiveness of mobile-health technology-based health behaviour change or disease management interventions for health care consumers: a systematic review. PLoS Med 2013;10(1):e1001362 [FREE Full text] [doi: 10.1371/journal.pmed.1001362] [Medline: 23349621]
28. Veritas Health Innovation. Covidence - Better Systematic Review Management. 2016. URL: https://www.covidence.org/home [accessed 2019-09-05]
29. Healthcare Information and Management Systems Society. 2014. HIMSS Patient Engagement Framework. URL: https://www.himss.org/himss-patient-engagement-framework [accessed 2019-07-13]
30. Walker DM, Hefner JL, Sieck CJ, Huerta TR, McAlearney AS. Framework for evaluating and implementing inpatient portals: a multi-stakeholder perspective. J Med Syst 2018 Jul 16;42(9):158. [doi: 10.1007/s10916-018-1009-3] [Medline: 30014267]
31. Athilingam P, Clochesy JM, Labrador MA. Intervention mapping approach in the design of an interactive mobile health application to improve self-care in heart failure. Comput Inform Nurs 2018 Feb;36(2):90-97. [doi: 10.1097/CIN.0000000000000383] [Medline: 28901967]
32. Lafata JE, Miller CA, Shires DA, Dyer K, Ratliff SM, Schreiber M. Patients' adoption of and feature access within electronic patient portals. Am J Manag Care 2018 Nov 1;24(11):e352-e357 [FREE Full text] [Medline: 30452203]
33. Ackerman SL, Sarkar U, Tieu L, Handley MA, Schillinger D, Hahn K, et al. Meaningful use in the safety net: a rapid ethnography of patient portal implementation at five community health centers in California. J Am Med Inform Assoc 2017 Sep 1;24(5):903-912 [FREE Full text] [doi: 10.1093/jamia/ocx015] [Medline: 28340229]
34. Raval MV, Taylor N, Piper K, Thakore M, Hoff K, Owens S, et al. Pediatric patient and caregiver preferences in the development of a mobile health application for management of surgical colorectal conditions. J Med Syst 2017 Jul;41(7):105. [doi: 10.1007/s10916-017-0750-3] [Medline: 28540617]
35. Krist AH, Peele E, Woolf SH, Rothemich SF, Loomis JF, Longo DR, et al. Designing a patient-centered personal health record to promote preventive care. BMC Med Inform Decis Mak 2011 Nov 24;11:73 [FREE Full text] [doi: 10.1186/1472-6947-11-73] [Medline: 22115059]

36. Shapiro-Mathews E, Barton AJ. Using the patient engagement framework to develop an institutional mobile health strategy. Clin Nurse Spec 2013;27(5):221-223. [doi: 10.1097/NUR.0b013e3182a0b9e2] [Medline: 23942098]
37. Anshari M, Almunawar M, Low PK, Wint Z, Younis MZ. Adopting customers' empowerment and social networks to encourage participations in e-health services. J Health Care Finance 2013;40(2):17-41. [Medline: 24551960]
38. Greysen SR, Harrison JD, Rareshide C, Magan Y, Seghal N, Rosenthal J, et al. A randomized controlled trial to improve engagement of hospitalized patients with their patient portals. J Am Med Inform Assoc 2018 Dec 1;25(12):1626-1633. [doi: 10.1093/jamia/ocy125] [Medline: 30346543]
39. van den Bulck SA, Hermens R, Slegers K, Vandenberghe B, Goderis G, Vankrunkelsven P. Designing a patient portal for patient-centered care: cross-sectional survey. J Med Internet Res 2018 Oct 1;20(10):e269 [FREE Full text] [doi: 10.2196/jmir.9497] [Medline: 30287416]
40. Wildenbos GA, Maasri K, Jaspers M, Peute L. Older adults using a patient portal: registration and experiences, one year after implementation. Digit Health 2018;4:2055207618797883 [FREE Full text] [doi: 10.1177/2055207618797883] [Medline: 30186619]
41. Wildenbos GA, Horenberg F, Jaspers M, Peute L, Sent D. How do patients value and prioritize patient portal functionalities and usage factors? A conjoint analysis study with chronically ill patients. BMC Med Inform Decis Mak 2018 Nov 21;18(1):108 [FREE Full text] [doi: 10.1186/s12911-018-0708-5] [Medline: 30463613]
42. Metting E, Schrage AJ, Kocks JW, Sanderman R, van der Molen T. Assessing the needs and perspectives of patients with asthma and chronic obstructive pulmonary disease on patient web portals: focus group study. JMIR Form Res 2018 Nov 22;2(2):e22 [FREE Full text] [doi: 10.2196/formative.8822] [Medline: 30684436]
43. Carman KL, Dardess P, Maurer M, Sofaer S, Adams K, Bechtel C, et al. Patient and family engagement: a framework for understanding the elements and developing interventions and policies. Health Aff (Millwood) 2013 Feb;32(2):223-231. [doi: 10.1377/hlthaff.2012.1133] [Medline: 23381514]
44. Health Quality Ontario. 2016. Engaging with Patients: Stories and Successes from the 2015/2016 Quality Improvement Plans. URL: http://www.hqontario.ca/Portals/0/documents/qi/qip/engaging-with-patients-en.pdf [accessed 2019-07-19]
45. Arkind J, Likumahuwa-Ackman S, Warren N, Dickerson K, Robbins L, Norman K, et al. Lessons learned from developing a patient engagement panel: an OCHIN report. J Am Board Fam Med 2015;28(5):632-638 [FREE Full text] [doi: 10.3122/jabfm.2015.05.150009] [Medline: 26355135]
46. Boyer AP, Fair AM, Joosten YA, Dolor RJ, Williams NA, Sherden L, et al. A multilevel approach to stakeholder engagement in the formulation of a clinical data research network. Med Care 2018 Oct;56(Suppl 10 Suppl 1):S22-S26 [FREE Full text] [doi: 10.1097/MLR.0000000000000778] [Medline: 30074947]
47. Hawke LD, Relihan J, Miller J, McCann E, Rong J, Darnay K, et al. Engaging youth in research planning, design and execution: practical recommendations for researchers. Health Expect 2018 Dec;21(6):944-949 [FREE Full text] [doi: 10.1111/hex.12795] [Medline: 29858526]
48. Hamilton CB, Hoens AM, Backman CL, McKinnon AM, McQuitty S, English K, et al. An empirically based conceptual framework for fostering meaningful patient engagement in research. Health Expect 2018 Feb;21(1):396-406 [FREE Full text] [doi: 10.1111/hex.12635] [Medline: 28984405]
49. Shelef DQ, Rand C, Streisand R, Horn IB, Yadav K, Stewart L, et al. Using stakeholder engagement to develop a patient-centered pediatric asthma intervention. J Allergy Clin Immunol 2016 Dec;138(6):1512-1517. [doi: 10.1016/j.jaci.2016.10.001] [Medline: 27744029]
50. Warren NT, Gaudino JA, Likumahuwa-Ackman S, Dickerson K, Robbins L, Norman K, et al. Building meaningful patient engagement in research: case study from ADVANCE clinical data research network. Med Care 2018 Oct;56(Suppl 10 Suppl 1):S58-S63 [FREE Full text] [doi: 10.1097/MLR.0000000000000791] [Medline: 30074953]
51. Chung AE, Vu MB, Myers K, Burris J, Kappelman MD. Crohn's and Colitis foundation of America partners patient-powered research network: patient perspectives on facilitators and barriers to building an impactful patient-powered research network. Med Care 2018 Oct;56(Suppl 10 Suppl 1):S33-S40 [FREE Full text] [doi: 10.1097/MLR.0000000000000771] [Medline: 30074949]
52. Faulkner M, Alikhaani J, Brown L, Cruz H, Davidson D, Gregoire K, et al. Exploring meaningful patient engagement in ADAPTABLE (aspirin dosing: a patient-centric trial assessing benefits and long-term effectiveness). Med Care 2018 Oct;56(Suppl 10 Suppl 1):S11-S15 [FREE Full text] [doi: 10.1097/MLR.0000000000000949] [Medline: 30074943]
53. Coathup V, Teare HJ, Minari J, Yoshizawa G, Kaye J, Takahashi MP, et al. Using digital technologies to engage with medical research: views of myotonic dystrophy patients in Japan. BMC Med Ethics 2016 Aug 24;17(1):51 [FREE Full text] [doi: 10.1186/s12910-016-0132-2] [Medline: 27553007]
54. Suarez-Balcazar Y. Empowerment and participatory evaluation of a community health intervention: implications for occupational therapy. OTJR (Thorofare N J) 2016 Jun 21;25(4):133-142. [doi: 10.1177/153944920502500402]
55. Canadian Institutes of Health Research. 2014. Strategy for Patient-Oriented Research - Patient Engagement Framework. URL: http://www.cihr-irsc.gc.ca/e/48413.html [accessed 2019-07-14]
56. Health Quality Ontario. 2016. Engaging With Patients and Caregivers About Quality Improvement: A Guide for Health Care Providers. URL: http://www.hqontario.ca/portals/0/documents/qi/qip/patient-engagement-guide-1611-en.pdf [accessed 2019-07-13]

57. Government of South Australia: SA Health. 2013. Guideline: Guide for Engaging with Consumers and the Community. URL: https://www.sahealth.sa.gov.au/wps/wcm/connect/f8d1d0004e454788aa0caf8ba24f3db9/Guideline_Guide+for+Engaging+with+Consumers+and+Community_Oct2015.pdf?MOD=AJPERES&CACHEID=f8d1d0004e454788aa0caf8ba24f3db9 [accessed 2019-07-14]
58. Hailey D. Institute of Health Economics. 2005. Consumer Involvement in Health Technology Assessment. URL: https://www.ihe.ca/download/consumer_involvement_in_health_technology_assessment.pdf [accessed 2019-07-14]
59. Health Consumers Queensland. 2017. Consumer and Community Engagement Framework. URL: http://www.hcq.org.au/our-work/framework/ [accessed 2019-07-14]
60. Hayes H, Buckland S, Tarpet M. INVOLVE. 2012. Briefing Notes For Researchers: Public Involvement in NHS, Public Health and Social Care Research. URL: http://www.invo.org.uk/wp-content/uploads/2014/11/9938_INVOLVE_Briefing_Notes_WEB.pdf [accessed 2019-07-14]
61. Health Quality Ontario. 2017. Ontario's Patient Engagement Framework: Creating a Strong Culture of Patient Engagement to Support High Quality Health Care. URL: https://www.hqontario.ca/Portals/0/documents/pe/ontario-patient-engagement-framework-en.pdf [accessed 2019-07-14]
62. Paediatric Integrated Cancer Services. 2013. A Toolkit for Consumer Participation and Engagement. URL: https://pics.org.au/health-professionals/engaging-consumers/ [accessed 2019-07-14]
63. James Lind Alliance. 2018. James Lind Alliance Guidebook. URL: http://www.jla.nihr.ac.uk/jla-guidebook/ [accessed 2019-07-14]
64. Vancouver Coastal Health. 2017. Community Engagement Framework. URL: http://www.vch.ca/Documents/CE-Framework.pdf [accessed 2019-07-14]
65. Center for Advancing Health. American Hospital Association. 2014. A New Definition of Patient Engagement: What is Engagement and Why is it Important? URL: http://www.hret-hiin.org/resources/display/a-new-definition-of-patient-engagement-what-is-engagement-and-why-is-it-important [accessed 2019-09-10]
66. Listowel Wingham Hospitals Alliance. 2017. Quality & Patient Engagement Framework. URL: http://www.lwha.ca/data/1/rec_docs/1294_LWHA_Quality_Framework_Final_2017.pdf [accessed 2019-07-14]
67. Agency for Clinical Innovation. 2015. Patient Experience and Consumer Engagement: A Framework for Action. URL: https://www.aci.health.nsw.gov.au/__data/assets/pdf_file/0005/256703/peace-framework.pdf [accessed 2019-07-14]
68. Canadian Agency for Drugs and Technologies in Health. 2019. CADTH Framework for Patient Engagement in Health Technology Assessment. URL: https://www.cadth.ca/cadth-framework-patient-engagement-health-technology-assessment [accessed 2019-07-13]
69. Health Quality Ontario. UW Departments Web Server. 2016. Choosing Methods for Patient and Caregiver Engagement. URL: https://depts.washington.edu/panport/resources/framework-guide/choosing-methods-patient-and-caregiver-engagement [accessed 2019-07-13]
70. Canadian Patient Safety Institute. 2018. Engaging Patients in Patient Safety – a Canadian Guide. URL: https://www.patientsafetyinstitute.ca/en/toolsResources/Patient-Engagement-in-Patient-Safety-Guide/Pages/default.aspx [accessed 2019-07-13]
71. NHS England. 2016. Patient Engagement in Patient Safety: A Framework for the NHS. URL: https://www.england.nhs.uk/signuptosafety/wp-content/uploads/sites/16/2016/05/pe-ps-framwrk-apr-16.pdf [accessed 2019-07-13]
72. Government of Prince Edward Island. 2016. Engagement Toolkit. URL: http://www.gov.pe.ca/photos/original/hpei_engagetool.pdf [accessed 2019-07-13]
73. Arsand E, Demiris G. User-centered methods for designing patient-centric self-help tools. Inform Health Soc Care 2008 Sep;33(3):158-169. [doi: 10.1080/17538150802457562] [Medline: 18850399]
74. Manafo E, Petermann L, Mason-Lai P, Vandall-Walker V. Patient engagement in Canada: a scoping review of the 'how' and 'what' of patient engagement in health research. Health Res Policy Syst 2018 Feb 7;16(1):5 [FREE Full text] [doi: 10.1186/s12961-018-0282-4] [Medline: 29415734]
75. Graffigna G, Barello S, Wiederhold B, Bosio AC, Riva G. Positive technology as a driver for health engagement. Stud Health Technol Inform 2013;191:9-17. [doi: 10.3233/978-1-61499-282-0-9] [Medline: 23792833]
76. Frampton SB, Guastello S, Hoy L, Naylor M, Sheridan S, Johnston-Fleece M. Harnessing evidence and experience to change culture: a guiding framework for patient and family engaged care. NAM Perspectives 2017 Jan 31;7(1):1-38. [doi: 10.31478/201701f]
77. University of Ottawa Heart Institute. 2018. Patient Engagement Framework. URL: https://www.ottawaheart.ca/patients-visitors/patient-engagement-framework [accessed 2019-07-13]
78. Abel EA, Shimada SL, Wang K, Ramsey C, Skanderson M, Erdos J, et al. Dual use of a patient portal and clinical video telehealth by veterans with mental health diagnoses: retrospective, cross-sectional analysis. J Med Internet Res 2018 Nov 7;20(11):e11350 [FREE Full text] [doi: 10.2196/11350] [Medline: 30404771]
79. Roblin DW, Houston 2nd TK, Allison JJ, Joski PJ, Becker ER. Disparities in use of a personal health record in a managed care organization. J Am Med Inform Assoc 2009;16(5):683-689 [FREE Full text] [doi: 10.1197/jamia.M3169] [Medline: 19567790]

80. Abelson J, Li K, Wilson G, Shields K, Schneider C, Boesveld S. Supporting quality public and patient engagement in health system organizations: development and usability testing of the Public and Patient Engagement Evaluation Tool. Health Expect 2016 Aug;19(4):817-827 [FREE Full text] [doi: 10.1111/hex.12378] [Medline: 26113295]
81. Esmail L, Moore E, Rein A. Evaluating patient and stakeholder engagement in research: moving from theory to practice. J Comp Eff Res 2015 Mar;4(2):133-145. [doi: 10.2217/cer.14.79] [Medline: 25825842]
82. Dukhanin V, Topazian R, DeCamp M. Metrics and evaluation tools for patient engagement in healthcare organization- and system-level decision-making: a systematic review. Int J Health Policy Manag 2018 May 16;7(10):889-903 [FREE Full text] [doi: 10.15171/ijhpm.2018.43] [Medline: 30316241]
83. Nilsen P. Making sense of implementation theories, models and frameworks. Implement Sci 2015 Apr 21;10:53 [FREE Full text] [doi: 10.1186/s13012-015-0242-0] [Medline: 25895742]

Abbreviations

EHR: electronic health record
IT: information technology
RQ: research question

Edited by G Eysenbach; submitted 13.05.19; peer-reviewed by T Risling, R Buie, K Fuji, D Alhuwail, R Lee; comments to author 18.06.19; revised version received 14.07.19; accepted 30.07.19; published 08.10.19.

Please cite as:
Leung K, Lu-McLean D, Kuziemsky C, Booth RG, Collins Rossetti S, Borycki E, Strudwick G
Using Patient and Family Engagement Strategies to Improve Outcomes of Health Information Technology Initiatives: Scoping Review
J Med Internet Res 2019;21(10):e14683
URL: https://www.jmir.org/2019/10/e14683
doi:10.2196/14683
PMID:31596241

Review

Barriers to the Use of Mobile Health in Improving Health Outcomes in Developing Countries: Systematic Review

Clemens Kruse[1], MHA, MSIT, MBA, PhD; Jose Betancourt[1], DPH; Stephanie Ortiz[1], BHA; Susana Melissa Valdes Luna[1], BA, MHA; Inderdeep Kaur Bamrah[1], BA, MHA; Narce Segovia[1], BA, MHA

School of Health Administration, Texas State University, San Marcos, TX, United States

Corresponding Author:
Clemens Kruse, MHA, MSIT, MBA, PhD
School of Health Administration
Texas State University
601 University Dr
Encino Hall 250
San Marcos, TX, 78666
United States
Phone: 1 210 355 4742
Fax: 1 512 245 8712
Email: scottkruse@txstate.edu

Abstract

Background: The use of mobile health (mHealth) technologies to improve population-level health outcomes around the world has surged in the last decade. Research supports the use of mHealth apps to improve health outcomes such as maternal and infant mortality, treatment adherence, immunization rates, and prevention of communicable diseases. However, developing countries face significant barriers to successfully implement, sustain, and expand mHealth initiatives to improve the health of vulnerable populations.

Objective: We aimed to identify and synthesize barriers to the use of mHealth technologies such as text messaging (short message service [SMS]), calls, and apps to change and, where possible, improve the health behaviors and health outcomes of populations in developing countries.

Methods: We followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses checklist. Deriving search criteria from the review's primary objective, we searched PubMed and CINAHL using an exhaustive terms search (eg, mHealth, text messaging, and developing countries, with their respective Medical Subject Headings) limited by publication date, English language, and full text. At least two authors thoroughly reviewed each article's abstract to verify the articles were germane to our objective. We then applied filters and conducted consensus meetings to confirm that the articles met the study criteria.

Results: Review of 2224 studies resulted in a final group of 30 articles for analysis. mHealth initiatives were used extensively worldwide for applications such as maternal health, prenatal care, infant care, HIV/AIDS prevention, treatment adherence, cardiovascular disease, diabetes, and health education. Studies were conducted in several developing countries in Africa, Asia, and Latin America. From each article, we recorded the specific health outcome that was improved, mHealth technology used, and barriers to the successful implementation of the intervention in a developing country. The most prominent health outcomes improved with mHealth were infectious diseases and maternal health, accounting for a combined 20/30 (67%) of the total studies in the analysis. The most frequent mHealth technology used was SMS, accounting for 18/30 (60%) of the studies. We identified 73 individual barriers and grouped them into 14 main categories. The top 3 barrier categories were infrastructure, lack of equipment, and technology gap, which together accounted for 28 individual barriers.

Conclusions: This systematic review shed light on the most prominent health outcomes that can be improved using mHealth technology interventions in developing countries. The barriers identified will provide leaders of future intervention projects a solid foundation for their design, thus increasing the chances for long-term success. We suggest that, to overcome the top three barriers, project leaders who wish to implement mHealth interventions must establish partnerships with local governments and nongovernmental organizations to secure funding, leadership, and the required infrastructure.

(J Med Internet Res 2019;21(10):e13263) doi:10.2196/13263

KEYWORDS

health outcomes; telemedicine; text messaging; communication barriers; developing countries; treatment outcome

Introduction

Background

Mobile devices are a cheap source of technology for addressing health care needs in developing countries. With the expansion of technology, mobile health (mHealth) is a tool that can be used to exchange health information for improving health outcomes through short message service (SMS) text messaging, mobile apps, and calls [1]. mHealth offers simplicity, efficiency, and effectiveness to patients due to its ability of rapid communication. mHealth intervention is a useful tool due to the ability to be accessible at the user's convenience. Mobile apps can be used to assess and measure the impact of a specific disease or may actually prevent a specific illness from occurring. A simple text can communicate, store, retrieve, and remind patients of their health status or deliver messages that promote healthy behaviors and choices. It is an inexpensive tool that can reduce the disparities of health in developing countries. Health care professionals now use smartphones or tablet computers to accomplish tasks for which they once used to need a pager or a personal digital assistant [2,3].

Definition of Key Terms

The term barrier is defined as "something that separates one thing from another" [4]. It is anything that prevents a certain goal from being achieved [4]. The second term, mHealth, is defined as a "medical and public health practice supported by mobile devices, such as mobile phones, patient monitoring devices, personal digital assistants, and other wireless devices" [5]. A developing country is a country that has a slow rate of industrialization, low per capita income, high unemployment, high poverty rate, and low standard of living [6]. Developing countries usually rely on developed countries for their economic growth and prosperity [6].

Rationale for the Review

mHealth is a tool that has had a positive impact on developed countries and has contributed to improving the health outcomes of populations around the world [7]. Specifically, researchers have focused on SMS in health care, and leading health organizations recommend its use [8]. Around the globe, mobile-cellular subscriptions will soon match the number of the population worldwide and are expected to continue to increase [9]. This is especially true in the developing world, where the market has not yet reached saturation [9]. mHealth has closed the gap in the digital divide in low-resource areas [10]. In the developing world, the World Health Organization reported a shortage of health care workers in 57 countries, resulting in a clear opportunity for innovative and effective solutions to help improve the health outcomes of their most vulnerable populations [11]. Mobile technology devices such as tablets, phones, computers, and tracking devices can be used to support and enhance health care in developing countries. The use of text messaging to promote healthy behaviors and healthy choices can be considered a groundbreaking component in improving population and community health.

Context of Other Evidence

Many studies have been carried out to determine the efficiency of mHealth in developing countries. A literature review of SMS-supported interventions for surveillance, management, treatment compliance, and prevention of noncommunicable diseases in India, South Africa, and Kenya found mobile phones to be well accepted by the population; however, high-quality intervention designed studies were needed [12]. In Nigeria, mobile device questionnaires were used to understand the perceptions of women at high risk of maternal death; although over 90% of women owned mobile phones, innovative methods were lacking to strengthen the delivery of maternal health information to those hard-to-reach populations [13]. In Zambia, SMS was found to have the potential to diagnose HIV early in infants by accelerating the delivery of results of blood sample testing to clinics, but the identification of lack of mobile phone ownership during the design of the study was found to hinder the success of the intervention [14].

In China, a smartphone app and text messaging were used to improve vaccination coverage among children, as well as the consumption of infant micronutrient powder packets [15]. Caregivers' suspicious beliefs and lack of acceptability were the major causes negatively affecting the success of the mHealth intervention [16,17]. In remote areas of Vietnam, the mMom app was used to improve pregnant women's maternal and infant health knowledge. An anticipated challenge was the high level of integration among local partners that required constant communication and engagement for coordination of the mHealth initiative [3].

An exploratory qualitative study conducted in Latin America and the Caribbean sought to further understand the needs of underserved populations and their exposure to public health interventions that used information and communications technologies to highlight the scarcity of such tools to reduce inequities. The greatest challenges were the lack of sustainability for financial and technical resources due to the unreliability of sustained external funding, poor intervention design caused by the resistance of precedents, and lack of technological literacy among participants unfamiliar with the use of information and communications technologies [18].

A study designed for primary prevention of hypertension in Argentina, Guatemala, and Peru found challenges that consisted of the unacceptability of mHealth innovations by the targeted communities, and emphasized the need to tailor the interventions to potential literacy challenges attributed to lack of understanding of cultural context [19]. In Brazil, a mobile phone–based intervention to promote prenatal care practices found that only one-fifth of women eligible for the study were actually interested in participating [19]. In Tajikistan, Bolivia, and Palestine, a behavioral change intervention was deployed using text messages to bring awareness of using contraceptives among the young to prevent unwanted pregnancies [20]. Negative attitudes toward and beliefs about contraception, including the cultural stigma of having sex before marriage,

being judged, and confidentiality concerns, limited participant discussion of contraception with providers [21]. In urban and rural areas of Guatemala, text messages were used to remind parents of infants to attend vaccination visits and decrease unnecessary morbidity [22]. This study concluded that client preference for delivery modalities such as a combination of text messaging and phone calls should have been considered to reach the maximum amount of the targeted population [22].

Objective

While other studies have focused on the potential benefits of mHealth in developing countries, our literature review sought to define the barriers that impede the successful application of mHealth (eg, SMS, cell phones, apps) interventions that aim to improve health outcomes of a population in diverse developing countries around the world. This knowledge may serve as a useful tool for project leaders to consider when planning or designing future mHealth interventions to strengthen the chances of long-term, sustainable success in the community. To analyze the use of mHealth in developing countries, more studies need to examine the different types of barriers these countries face. We conducted this literature review to determine what type of mHealth initiatives are more popular in developing countries, as well as the outcomes and barriers identified by the respective article authors. We aimed to provide a clearer understanding of what initiatives have the best supporting evidence of improving health outcomes by using mHealth approaches and of the resources developing countries require to foster the long-term and sustainable success of these projects.

Methods

Protocol Registration and Eligibility Criteria

This review followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Multimedia Appendix 1) [23]. We did not register the review. The main objective was to identify and synthesize the barriers to the use of mHealth to improve the health outcomes in developing countries. Articles were eligible for review if they met criteria such as having a health outcome and involving the use of mHealth technology in a developing country. We analyzed articles only if (1) the full-text article was available, (2) the article related to humans, (3) the article was published between 2008 and 2018, and (4) the article was written in English. Exclusion criteria for this review were systematic reviews, articles unrelated to the objective of the review, and no direct health outcome being involved. We did not consider for this review any studies that were in progress. Finally, we removed duplicate articles from the literature matrix.

Information Sources and Search

In addition to reporting this review in accordance with the PRISMA guidelines, we conducted it using techniques of the Assessment of Multiple Systematic Reviews [24]. We searched PubMed and CINAHL with an exhaustive search string comprising Medical Subject Headings (MeSH), due to their widespread availability, and Boolean operators. Multimedia Appendix 2 provides a detailed list of MeSH terms. We conducted searches for this literature review in September 2018.

Risk of Bias in Individual Studies

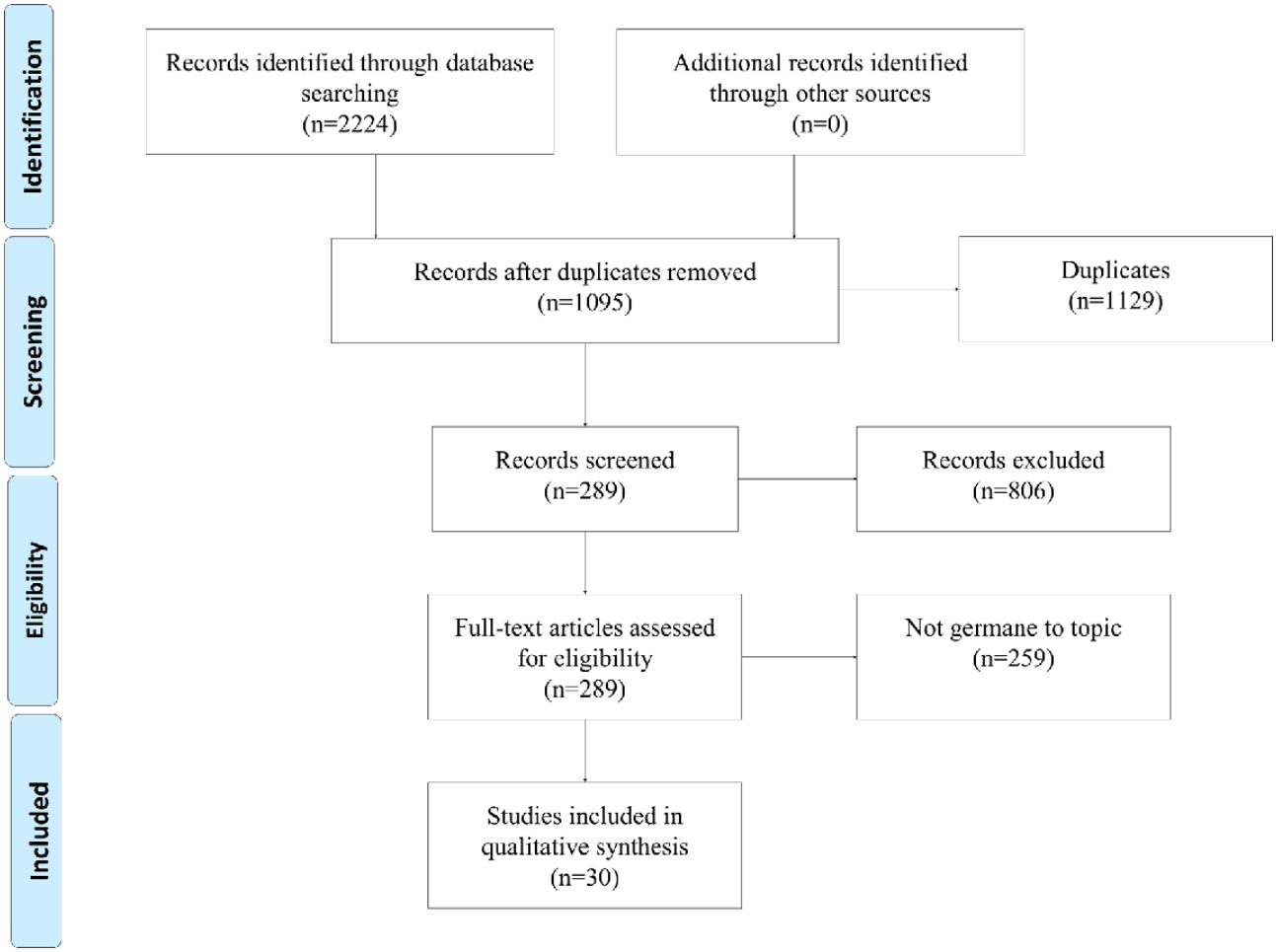
Our methods did not enable randomization, so to control for selection bias, we conducted the search using exhaustive MeSH terms, which are widely used in databases used for literature research. To avoid influencing each other's individual opinions, each researcher recorded his or her findings independently. We calculated a kappa statistic to measure interrater reliability, which refers to the agreement and consistency of article selection. The kappa was .78, which reflects a moderate agreement between the reviewers [25,26].

Results

Study Selection, Data Collection Process, and Data Items

Figure 1 illustrates the search and selection process in each database. The initial search in PubMed resulted in 535 articles, and in CINAHL the search produced 1689 articles. The total number of articles obtained from both search engines was 2224. Then, after applying limiters, we narrowed the PubMed results down to 144 articles and CINAHL to 145 articles. All 289 abstracts were screened by at least two reviewers. Reviewers screened abstracts recording their individual observations and recommendations to either include or exclude each article. A consensus meeting was then held to arrive at an agreement on the selected articles. If there was a disagreement, a third author's vote was required to reach consensus. We produced a list of 58 abstracts germane to the objective of the review. We distributed these articles among our team in a manner that ensured each article was analyzed by at least two reviewers. A second consensus meeting arrived at a final group of articles for analysis of 30. The other articles were eliminated after a full reading.

Figure 1. Preferred Reporting Items for Systematic Reviews and Meta-Analyses flow diagram of the literature search and selection process.



Study Selection and Characteristics

Between September and November 2018, we reviewed the 30 articles germane to this review's objective of defining the barriers to implementing mHealth interventions in developing countries. Results included mHealth initiatives covering health outcomes such as prenatal care, infectious diseases, medication adherence, appointment reminders, and chronic diseases education. The articles in the group for analysis reported interventions piloted across several developing countries around the world. We divided the 30 articles among the reviewers for analysis by at least two reviewers. We held group consensus meetings to facilitate discussion of individual barriers identified in the articles.

Results of Individual Studies and Synthesis of Results

We analyzed the articles to determine the health outcomes most commonly improved with mHealth interventions, the most used mHealth technology, and, most importantly, the barriers that hindered the adoption or impact of the mHealth interventions. We break down each of the findings to provide a comprehensive vision for future leaders who wish to implement mHealth interventions in developing countries.

Among the articles analyzed, maternal health was the most prevalent health outcome with a frequency of 9 out of the 30 articles (30%) [27-35]. Infectious diseases and chronic diseases were the second most prevalent health outcomes, with a frequency of each of 8 of 30 (27%) articles [36,39-44,49-51]. The third most prevalent outcome was preventive health, occurring in 5 of 30 (17%) articles [22,52-55].

The mHealth interventions identified in the literature were SMS only; SMS or phone calls and voice messages, or both SMS and phone call and voice messages; SMS phone app; multimedia messages for diagnosis; and a combination of SMS, smartphone app, and cellphones. We identified SMS in 28 of the 30 articles we analyzed (93%). The use of SMS only was mentioned in 16 of 26 interventions (62%) [28-31,35,37,38,40-43,45-47,50,52]. SMS with or without phone calls and voice messages was the second highest with a frequency of 4 of 26 interventions (15%) [22,32,39,53]. Smartphone with the use of apps was the third most commonly mHealth used with a frequency of 3 of 26 interventions (12%) [34,36,55]. Interventions using multimedia messages had a frequency of 2 of 26 (4%) [27,54]. Finally, the combination of SMS, smartphone app, and cell calls occurred in only 1 of 26 (4%) interventions [33].

We classified the 73 barriers into 14 categories identified in the 30 articles. The categories are immature (or lack of) infrastructure (10/73, 14%), lack of equipment (9/73, 12%), technology gap (9/73, 12%), human resource issues (7/73, 10%), time or work conflict (7/73, 10%), cost (6/73, 8%), lack of public policy (8/73, 8%), literacy (4/73, 5%), language barriers (4/73, 5%), psychosocial issues (4/73, 5%) lack of training (3/73, 4%), concerns about privacy and confidentiality of information (2/73,

3%), lack of efficacy (1/73, 1%), and exposure of program (1/73, 1%).

Figure 2 provides a geographic distribution of the studies in the analyzed studies. Africa accounted for 16 of 30 (53%) articles, Asia for 10 of 30 (33%) articles, and South America for 3 of 30 (10%) articles; 1 article studied both Africa and South America (3%).

Table 1 summarizes the following details of each article analyzed: study design, sample size, technological intervention, health outcome, barriers identified, and world region.

Figure 2. Location of studies of mobile health technologies to improve preventive health outcomes in developing countries, by region.

Table 1. Summary of results.

First author, date, reference	Study design and sample size	mHealth intervention category	Health outcome category	Barriers identified	Continent
Ginsburg, 2015 [36]	Open data-kit survey (design-stage evaluation activity)	Smartphone app	Infectious disease	• Technology gap • Language barrier	Africa
Nhavoto, 2017 [37]	RCT[a] and interviews (tuberculosis n=69, HIV n=72)	SMS[b] only	Infectious disease	• Privacy concerns • Literacy • Language barrier • Equipment	Africa
Bediang, 2014 [38]	Blinded RCT (intervention n=104, control n=104)	SMS only	Infectious disease	• Infrastructure • Cost • Policy	Africa
Bigna, 2013 [39]	RCT (n=224 divided into 4 groups)	SMS with or without phone calls and voice mail	Infectious disease	• Equipment • Language barrier • Policy • Privacy concerns • Infrastructure	Africa
Medhanyie, 2015 [27]	Interviews, 2893 electronic health records of 1122 women	Multimedia messages for diagnosis	Maternal health	• Time or work conflict • Human resources issues • Time or work conflict • Time or work conflict	Africa
Rokicki, 2017 [28]	Cluster RCT (n=756)	SMS only	Maternal health	• Time or work conflict • Equipment • Cost	Africa
Toda, 2016 [52]	Clustered RCT (intervention n=32 [88 cases], control n=32 [21 cases])	SMS only	Preventive health	• Human resources issues • Technology gap • Training	Africa
Flax, 2017 [29]	Cluster RCT and interviews (n=195)	SMS only	Maternal health	• Infrastructure • Psychosocial stressors	Africa
Ngabo, 2012 [30]	Pilot study	SMS only	Maternal health	• Infrastructure • Equipment • Cost	Africa
Jia, 2015 [53]	Longitudinal data analysis of monthly electronic health record suspect and mortality cases for both the traditional sentinel program and mobile phone reporting (n=178)	SMS with or without phone calls and voice mail	Preventive health	• Training	Africa
Leon, 2015 [40]	RCT (n=22 studied, n=15 interviewed)	SMS only	Chronic disease	• Technology gap • Infrastructure • Psychosocial stressors	Africa
Hao, 2015 [41]	Interviews (n=11)	SMS only	Infectious disease	• Time or work conflict • Equipment • Policy	Africa
Lund, 2014 [31]	Open-label, pragmatic-cluster RCT, 2550 pregnant women (intervention n=1311, control n=1239)	SMS only	Maternal health	• Equipment • Literacy	Africa
Tuijn, 2011 [54]	Feasibility study attaching a cell phone to a microscope	Multimedia messages for diagnosis	Preventive health	• Infrastructure • Equipment • Training	Africa

First author, date, reference	Study design and sample size	mHealth intervention category	Health outcome category	Barriers identified	Continent
Linnemayr, 2017 [42]	True experiment (n=332)	SMS only	Infectious disease	• Technology gap	Africa
Steury, 2016 [43]	RCT (n=96)	SMS only	Chronic disease	• Efficacy • Cost	Africa
Ippoliti, 20107 [32]	RCT (n=17)	SMS with or without phone calls and voice mail	Maternal health	• Infrastructure • Literacy	Africa and South America
Uddin, 2017 [33]	Stratified 2-stage random-cluster sampling technique to select participants (n=5280)	SMS, smartphone app, and cell phone calls	Maternal health	• Technology gap	Asia
Balakrishnan, 2016 [34]	Feasibility study (n=512)	Smartphone app	Maternal health	• Cost • Training	Asia
Fang, 2018 [44]	RCT (n=247)	SMS with or without phone calls and voice mail	Chronic disease	• Technology gap • Equipment • Psychosocial stressors	Asia
Zhou, 2016 [45]	RCT and the intervention period lasted for 6 months and used baseline and follow-up surveys (n=351 villages, n=1393 children)	SMS only	Chronic disease	• Policy	Asia
Fang, 2016 [46]	Random sampling method RCT (3 groups: SMS [n=95] SMS + Micro Letter app [n=92], and phone [n=93])	SMS only	Chronic disease	• Technology gap • Infrastructure	Asia
Kumboyono, 2017 [47]	No RCT, posttest-only control group design (n=90)	SMS only	Infectious disease	• Psychosocial stressors	Asia
Kazi, 2013 [48]	No RCT, systematic random sampling (n=28)	SMS only	Infectious disease	• Human resources issues	Asia
Mohan, 2017 [49]	Cross-sectional study (n=800)	SMS only	Chronic disease	• Language barrier	Asia
Lin, 2017 [35]	True experiment (n=757)	SMS only	Maternal health	• Training • Policy • Exposure of program • Time or work conflict	Asia
Wu, 2014 [55]	Prospective true experiment (intervention n=97, control n=128)	Smartphone app	Preventive health	• N/A[c]	Asia
Beratarrechea, 2015 [22]	Interviews (n=43)	SMS with or without phone calls and voice mail	Preventive health	• Infrastructure • Technology gap	South America
Rico, 2017 [50]	Interviews (n=14)	SMS only	Chronic disease	• Literacy • Technology gap	South America
Piette, 2012 [51]	RCT (n=200)	SMS only	Chronic disease	• Cost • Equipment • Infrastructure • Policy	South America

[a]RCT: randomized controlled trial.

[b]SMS: short message service.

[c]N/A: not available.

Discussion

Principal Findings

This review identified the common barriers faced by developing countries in the adoption of mHealth. mHealth is widely used in developing countries as a tool to improve the health outcomes of highly vulnerable communities and individuals. Based on the evidence found in this review, mHealth is an effective method to support health care services. mHealth has been used in many developing countries in regions such as Africa, Asia, and Latin America. These countries constantly battle infectious diseases, chronic diseases, perinatal complications, acute diseases, birth defects, and many more. This review revealed important barriers that must be understood before implementing mHealth initiatives. Considering and assessing these barriers prior to the design phase of an mHealth intervention will have a positive impact on the health outcomes of populations and individuals.

As noted above, mHealth can provide a great opportunity to solve health care issues faced by developing countries. However, there are various challenges and barriers to be considered prior to implementation.

Health Outcomes

Based on the findings of this review, the 2 main health outcomes affected by an mHealth intervention were infectious diseases and maternal health. In developing countries, the burden of infectious diseases is prevalent due to poverty, leading to "poor nutrition, indoor air pollution and lack of access to proper sanitation, and lack of health education" [56]. According to the World Health Organization, most illnesses are avoidable and treatable. It is estimated that diseases account for up to 45% of the burden in poor countries due to poverty; HIV, tuberculosis, and malaria account for 18% [56]. The second most common health outcome affected by an mHealth intervention was maternal health. Developing countries account for 99% of all maternal deaths compared with developed countries [53]. This is the biggest health gap in the world [53]. In remote locations, poor women are more prone to receive inadequate care, specifically in the areas lacking skilled health care workers [27].

The majority of the reviewed articles used SMS as an mHealth intervention to improve infectious disease, health outcomes, and patient treatment adherence. In developing countries, infectious diseases are prevalent due to the lack of preventive care. mHealth interventions in developing countries were considered effective in improving antenatal care, vaccination, and preventive treatment for chronic and infectious diseases [55]. SMS was also effective for maternal health, prenatal care, infant care, HIV/AIDS prevention, treatment adherence, cardiovascular disease care, diabetes care, health education, tuberculosis prevention and care, anemia care, immunization, and disease awareness [32,49]. SMS directly increased disease awareness by providing health tips and reinforcing reminder systems. Moreover, SMS provided emotional support to patients, promoted knowledge about health, and influenced attitude change toward greater self-responsibility [36].

Mobile Health Tool Used

SMS was the most commonly used mHealth tool due to the number of mobile phones in use. An estimated 4.5 billion people are expected to have mobile phones worldwide by 2020 [33]. Compared with other methods of communication, text messaging has an advantage due to its low cost and high reliability [33]. Researchers in the field state that educational mHealth training programs are effective in raising awareness by offering an efficient and cost-effective way to achieve the success of mHealth implementation [42]. A health education approach via mobile phones can be used to manage diseases, aid medical testing, and improve treatments. Specifically, text messaging can be used for interventions and health education, because it is particularly popular in developing countries [42]. Through simple text messaging, patients have reported that they felt more confident in their treatment [49]. SMS has the potential to make patients feel supported, encouraged, and aware, thus helping them take better care of themselves and continue treatment [28]. SMS reminders also improve appointment attendance and SMS text messaging helps health care providers prescribe medicine on a timely basis, consequently improving patient care [32,39].

SMS is perceived as a tool that can boost the rate of adherence to medical treatment and has the potential to help in the prevention of diseases. Health information transmitted through text messages can also effectively be used to manage the treatment of infectious diseases such as tuberculosis [45]. The main mechanism is the use of text messages to remind patients about appointments and taking medications, to deliver motivational messages and health education or health promotion messages. Another approach is the use of mobile phones by health workers to help support services in diagnosing women and children in remote areas and identifying patients at risk who need to be referred. Throughout the literature review, SMS was often mentioned in combination with phone calls, voice messages, smartphone apps, and multimedia messages. However, the literature also provided examples of participants who did not know how to use mobile phones or similar technology efficiently. In other words, they did not know how to work a phone or were unable to read text messages. This limited knowledge may be attributed to the amount of exposure individuals in developing countries have to technology. Therefore, there was a correlation between frequency of phone usage and knowledge of this technology. Phone calls as an mHealth intervention can also help improve health outcomes while at the same time offering participants a simpler method than SMS. Mobile phone apps were proven to be an efficient tool to assist people in achieving early screening for diagnosis and treatment purposes. Moreover, apps have been shown to prevent health complications, thus helping improve preventive medicine [50].

World Region

Africa was the most frequent study setting in the articles we analyzed. This review suggested that this region has been extensively participating in mHealth projects.

The 3 most prevalent barrier categories were lack of infrastructure, lack of equipment, and technology gap.

Developing countries should consider investing in their infrastructure and encouraging partnerships with equipment providers to help their populations afford phones and learn how to use them. It is vital for developing countries to adapt to new emerging technologies in an effort to reduce the risk of being left behind in the great technological advancements in health.

Strengths and Limitations

This review had several limitations. First, selection bias tended to be prevalent in many research studies. To help address this problem, we held consensus meetings once per week to discuss the findings of the research articles. However, our main controls for selection bias were (1) identifying the research objective, (2) defining the key terms used, and (3) having more than one reviewer examine each article. We conducted all consensus meetings either through Skype or in person. These meetings offered great value to our research because they reduced personal bias when eliminating the articles from the literature matrix. We gathered feedback, opinions, and knowledge throughout the process. Another selection bias was the selection of only free full-text articles. We eliminated a few articles in this step. Including those articles in the review most likely would not have changed the outcome of our review, but it might have identified additional barriers.

Second, we examined only 10 years' worth of articles when abstracting the data. However, this may or may not be a limitation, as the technology used in mHealth may not have existed earlier. A third possible limitation is publication bias for the 10 years considered in this review.

This review adds to the body of knowledge on the significant barriers mHealth confronts in developing countries. This review was constructed in accordance with PRISMA guidelines. We limited our review to 2 well-known research databases, CINAHL and PubMed. As a result, we expect this review will have a high external and internal validity.

Recommendations and Suggestions for Further Research

To overcome these barriers, the published literature suggested some important solutions. Strengthening health care systems through the use of mHealth requires strong governance, as well as the commitment of the private sector [35,48]. More investment in phones and rigorous training on these devices is also required to improve their acceptability in developing countries [47]. It is also important to consider the characteristics of the population, such as socioeconomic background, to gain a better perspective of the community [29]. When phone ownership is lacking, a microcredit program, in which several people can obtain a loan to purchase a group phone, may be feasible, and consequently the group would rely on other family members or the community to improve their health [37]. The successful development of interventions using the capability of mHealth technologies lies within the criticality of mHealth research. It entails important characteristics, such as collaboration throughout all phases of the project [43].

It is important to adapt and redesign emerging interventions as the technology advances. The future of mHealth in both developed and developing countries is expected to be prosperous with new innovations arising exponentially throughout the health care domain. It is important to assess the disparities by country in order to improve their respective health care sectors. Community needs could be addressed and improved through the use of available technology by country. However, to drastically make a change and improve the use of mHealth in developing countries, policy reform at all levels is needed.

Project leader support through policy reform could compensate for the barriers faced in developing countries. Therefore, there is a need for future research on how governments can help their countries reach their goals to improve and increase the acceptance of mHealth as a means to improve health care and, ultimately, improve the health of their communities. In addition, there is a need for future implementation of mHealth technologies such as text messaging to improve chronic diseases, such as tuberculosis, HIV, hypertension, cardiovascular disease, colorectal cancer, and pneumonia, in remote and resource-limited settings to overcome the challenges a community faces. Implementation of mHealth initiatives requires rigorous training of health care workers, as well as of the designated population who will be participants in studies, to understand and use the technology correctly [39].

Training on the use of devices, such as cell phones and mobile apps, and on sharing and receiving text messages will not only improve the performance but also increase the acceptability of mHealth within the community [39,46]. Special attention needs to be paid to the illiterate when using SMS due to the inability of participants to read and comprehend these messages [55]. Lastly, there is a need to design the health system based on approaches to control the timing of text messaging, mobile network fluctuations, and mobile phone turnovers to improve treatment adherence and follow-up visits in cases of chronic diseases, infectious diseases, maternal care, and birth defects [32,48].

Conclusion

The published literature demonstrates the barriers faced by developing countries in the use of mHealth to improve health outcomes. This systematic review shed light on the most prominent health outcomes that can be improved using mHealth technology interventions in developing countries. SMS technology is readily available at low cost in developing countries and can be easily adopted to interventions that improve the health outcomes already identified. Additionally, the barriers identified will provide the leaders of future intervention projects a solid foundation for the design of those interventions, thus increasing the chances of long-term success and sustainability. We suggest that, to overcome the top 3 barriers, project leaders who wish to implement mHealth interventions must establish partnerships with local governments and nongovernmental organizations to secure funding, leadership, and the required infrastructure. This research identified the barriers and the frequency of those barriers by region. It also identified the most used type of mHealth tool, as well as the health outcomes affected by the tool used. This literature review highlighted the need for policy reform in developing countries to improve health care and, ultimately, improve the health of their communities.

Acknowledgments

The authors received no financial support for the research, authorship, or publication of this paper.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) checklist.
[PDF File (Adobe PDF File), 67 KB - jmir_v21i10e13263_app1.pdf]

Multimedia Appendix 2
Key terms used in the search string.
[PDF File (Adobe PDF File), 22 KB - jmir_v21i10e13263_app2.pdf]

References

1. Pillay Y, Motsoaledi PA. Digital health in South Africa: innovating to improve health. BMJ Glob Health 2018;3(Suppl 2):e000722 [FREE Full text] [doi: 10.1136/bmjgh-2018-000722] [Medline: 29713513]
2. Seebregts C, Fogwill T, Tanna G, Barron P, Benjamin P. MomConnect: an exemplar implementation of the health normative standards framework in South Africa. South Afr Health Rev 2016 Jan 01;2016(1):125-135 [FREE Full text]
3. McBride B, Nguyen LT, Wiljer D, Vu NC, Nguyen CK, O'Neil J. Development of a maternal, newborn and child mHealth intervention in Thai Nguyen Province, Vietnam: protocol for the mMom Project. JMIR Res Protoc 2018 Jan 11;7(1):e6 [FREE Full text] [doi: 10.2196/resprot.7912] [Medline: 29326095]
4. Odetola TD, Ayamolowo LB, Ayamolowo SJ. Childbearing women's perception about the use of mHealth for maternal health information in rural communities, Ile-Ife, Nigeria. J Int Soc Telemed eHealth 2018 May 08;6(1):9-1 [FREE Full text] [doi: 10.29086/JISfTeH.6.e9]
5. Kay M, Santos J, Takane M. mHealth: new horizons for health through mobile technologies. Global Observatory for eHealth series. Volume 3. Geneva, Switzerland: World Health Organization; 2011. URL: https://apps.who.int/iris/bitstream/handle/10665/44607/9789241564250_eng.pdf;jsessionid=FB42680AE27DBD769D572E8EC3ED937C?sequence=1 [accessed 2019-08-16]
6. Uddin MJ, Shamsuzzaman M, Horng L, Labrique A, Vasudevan L, Zeller K, et al. Use of mobile phones for improving vaccination coverage among children living in rural hard-to-reach areas and urban streets of Bangladesh. Vaccine 2016 Jan 04;34(2):276-283 [FREE Full text] [doi: 10.1016/j.vaccine.2015.11.024] [Medline: 26647290]
7. Domek GJ, Contreras-Roldan IL, Asturias EJ, Bronsert M, Bolaños Ventura GA, O'Leary ST, et al. Characteristics of mobile phone access and usage in rural and urban Guatemala: assessing feasibility of text message reminders to increase childhood immunizations. Mhealth 2018;4:9 [FREE Full text] [doi: 10.21037/mhealth.2018.03.05] [Medline: 29780812]
8. Wang X, Luo R, Liu C, Zhang L, Yue A, Medina A, et al. Using daily text messages to improve adherence to infant micronutrient powder (MNP) packets in rural western China: a cluster-randomized controlled trial. PLoS ONE 2018 Jan 19;13(1):e0191549. [doi: 10.1371/journal.pone.0191549] [Medline: 29352304]
9. Sutcliffe CG, Thuma PE, van Dijk JH, Sinywimaanzi K, Mweetwa S, Hamahuwa M, et al. Use of mobile phones and text messaging to decrease the turnaround time for early infant HIV diagnosis and notification in rural Zambia: an observational study. BMC Pediatr 2017 Mar 08;17(1):66 [FREE Full text] [doi: 10.1186/s12887-017-0822-z] [Medline: 28270134]
10. Krishnan A, Ferro EG, Weikum D, Vagenas P, Lama JR, Sanchez J, et al. Communication technology use and mHealth acceptance among HIV-infected men who have sex with men in Peru: implications for HIV prevention and treatment. AIDS Care 2015;27(3):273-282 [FREE Full text] [doi: 10.1080/09540121.2014.963014] [Medline: 25285464]
11. Albino S, Tabb KM, Requena D, Egoavil M, Pineros-Leano MF, Zunt JR, et al. Perceptions and acceptability of short message services technology to improve treatment adherence amongst tuberculosis patients in Peru: a Focus Group Study. PLoS One 2014;9(5):e95770 [FREE Full text] [doi: 10.1371/journal.pone.0095770] [Medline: 24828031]
12. Snyders FJ. Determining the Feasibility of Using Mobile Phones to Strengthen the Information Management of Preventative Health Care in South Africa [master's thesis]. Stellenbosch, South Africa: Stellenbosch University; 2013.
13. World Health Organization. Maternal mortality. Geneva, Switzerland: WHO; 2018. URL: https://www.who.int/en/news-room/fact-sheets/detail/maternal-mortality [accessed 2018-12-01] [WebCite Cache ID 74Knn43Yc]
14. Déglise C, Suggs LS, Odermatt P. SMS for disease control in developing countries: a systematic review of mobile health applications. J Telemed Telecare 2012;18(5):273-281. [doi: 10.1258/jtt.2012.110810] [Medline: 22826375]

15. Wu Q, Zhang Y, Chang S, Wang W, van Velthoven MH, Han H, et al. Monitoring and evaluating the adherence to a complementary food supplement (Ying Yang Bao) among young children in rural Qinghai, China: a mixed methods evaluation study. J Glob Health 2017 Jun;7(1) [FREE Full text] [doi: 10.7189/jogh.07.011101] [Medline: 28702176]
16. Farach N, Faba G, Julian S, Mejía F, Cabieses B, D'Agostino M, et al. Stories from the field: the use of information and communication technologies to address the health needs of underserved populations in Latin America and the Caribbean. JMIR Public Health Surveill 2015;1(1):e1 [FREE Full text] [doi: 10.2196/publichealth.4108] [Medline: 27227124]
17. Oliveira-Ciabati L, Vieira CS, Franzon ACA, Alves D, Zaratini FS, Braga GC, et al. PRENACEL - a mHealth messaging system to complement antenatal care: a cluster randomized trial. Reprod Health 2017 Nov 07;14(1):146 [FREE Full text] [doi: 10.1186/s12978-017-0407-1] [Medline: 29116028]
18. Rajan JV, Moura J, Gourley G, Kiso K, Sizilio A, Cortez AM, et al. Understanding the barriers to successful adoption and use of a mobile health information system in a community health center in São Paulo, Brazil: a cohort study. BMC Med Inform Decis Mak 2016 Nov 17;16(1):146 [FREE Full text] [doi: 10.1186/s12911-016-0385-1] [Medline: 27855685]
19. Vieira EM, Vieira CS, Bonifácio LP, de Oliveira Ciabati LM, Franzon AC, Zaratini FS, et al. PRENACEL: development and evaluation of an m-health strategy to improve prenatal care in Brazil. Health 2016 Aug 01;15:16. [doi: 10.5281/zenodo.1126866]
20. McCarthy OL, Wazwaz O, Osorio Calderon V, Jado I, Saibov S, Stavridis A, et al. Development of an intervention delivered by mobile phone aimed at decreasing unintended pregnancy among young people in three lower middle income countries. BMC Public Health 2018 May 02;18(1):576 [FREE Full text] [doi: 10.1186/s12889-018-5477-7] [Medline: 29716571]
21. Diez-Canseco F, Toyama M, Ipince A, Perez-Leon S, Cavero V, Araya R, et al. Integration of a technology-based mental health screening program into routine practices of primary health care services in Peru (the Allillanchu Project): development and implementation. J Med Internet Res 2018 Mar 15;20(3):e100 [FREE Full text] [doi: 10.2196/jmir.9208] [Medline: 29588272]
22. Beratarrechea A, Diez-Canseco F, Fernández A, Kanter R, Letona P, Martinez H, et al. [Acceptability of a mobile health based intervention to modify lifestyles in prehypertensive patients in Argentina, Guatemala and Peru: a pilot study]. Rev Peru Med Exp Salud Publica 2015;32(2):221-229 [FREE Full text] [Medline: 26338378]
23. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Int J Surg 2010;8(5):336-341 [FREE Full text] [doi: 10.1016/j.ijsu.2010.02.007] [Medline: 20171303]
24. Shea BJ, Grimshaw JM, Wells GA, Boers M, Andersson N, Hamel C, et al. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. BMC Med Res Methodol 2007 Feb 15;7:10 [FREE Full text] [doi: 10.1186/1471-2288-7-10] [Medline: 17302989]
25. Light RJ. Measures of response agreement for qualitative data: Some generalizations and alternatives. Psychol Bull 1971;76(5):365-377. [doi: 10.1037/h0031643]
26. McHugh ML. Interrater reliability: the kappa statistic. Biochem Med (Zagreb) 2012;22(3):276-282 [FREE Full text] [Medline: 23092060]
27. Medhanyie AA, Little A, Yebyo H, Spigt M, Tadesse K, Blanco R, et al. Health workers' experiences, barriers, preferences and motivating factors in using mHealth forms in Ethiopia. Hum Resour Health 2015 Jan 15;13:2 [FREE Full text] [doi: 10.1186/1478-4491-13-2] [Medline: 25588973]
28. Rokicki S, Cohen J, Salomon JA, Fink G. Impact of a text-messaging program on adolescent reproductive health: a cluster-randomized trial in Ghana. Am J Public Health 2017 Feb;107(2):298-305. [doi: 10.2105/AJPH.2016.303562] [Medline: 27997236]
29. Flax VL, Ibrahim AU, Negerie M, Yakubu D, Leatherman S, Bentley ME. Group cell phones are feasible and acceptable for promoting optimal breastfeeding practices in a women's microcredit program in Nigeria. Matern Child Nutr 2017 Jan;13(1):1 [FREE Full text] [doi: 10.1111/mcn.12261] [Medline: 26990786]
30. Ngabo F, Nguimfack J, Nwaigwe F, Mugeni C, Muhoza D, Wilson DR, et al. Designing and implementing an innovative SMS-based alert system (RapidSMS-MCH) to monitor pregnancy and reduce maternal and child deaths in Rwanda. Pan Afr Med J 2012;13:31 [FREE Full text] [Medline: 23330022]
31. Lund S, Nielsen BB, Hemed M, Boas IM, Said A, Said K, et al. Mobile phones improve antenatal care attendance in Zanzibar: a cluster randomized controlled trial. BMC Pregnancy Childbirth 2014 Jan 17;14:29 [FREE Full text] [doi: 10.1186/1471-2393-14-29] [Medline: 24438517]
32. Ippoliti NB, L'Engle K. Meet us on the phone: mobile phone programs for adolescent sexual and reproductive health in low-to-middle income countries. Reprod Health 2017 Jan 17;14(1):11 [FREE Full text] [doi: 10.1186/s12978-016-0276-z] [Medline: 28095855]
33. Uddin J, Biswas T, Adhikary G, Ali W, Alam N, Palit R, et al. Impact of mobile phone-based technology to improve health, population and nutrition services in Rural Bangladesh: a study protocol. BMC Med Inform Decis Mak 2017 Jul 06;17(1):101 [FREE Full text] [doi: 10.1186/s12911-017-0502-9] [Medline: 28683742]
34. Balakrishnan R, Gopichandran V, Chaturvedi S, Chatterjee R, Mahapatra T, Chaudhuri I. Continuum of care services for maternal and child health using mobile technology–a health system strengthening strategy in low and middle income

countries. BMC Med Inform Decis Mak 2016 Jul 07;16:84 [FREE Full text] [doi: 10.1186/s12911-016-0326-z] [Medline: 27387548]
35. Lin Q, Yang L, Li F, Qin H, Li M, Chen J, et al. A village‐based intervention: promoting folic acid use among rural Chinese women. Nutrients 2017 Feb 21;9(2):1 [FREE Full text] [doi: 10.3390/nu9020174] [Medline: 28230798]
36. Ginsburg AS, Delarosa J, Brunette W, Levari S, Sundt M, Larson C, et al. mPneumonia: development of an innovative mHealth application for diagnosing and treating childhood pneumonia and other childhood illnesses in low-resource settings. PLoS One 2015;10(10):e0139625 [FREE Full text] [doi: 10.1371/journal.pone.0139625] [Medline: 26474321]
37. Nhavoto JA, Grönlund Å, Klein GO. Mobile health treatment support intervention for HIV and tuberculosis in Mozambique: Perspectives of patients and healthcare workers. PLoS One 2017;12(4):e0176051 [FREE Full text] [doi: 10.1371/journal.pone.0176051] [Medline: 28419149]
38. Bediang G, Stoll B, Elia N, Abena J, Nolna D, Chastonay P, et al. SMS reminders to improve the tuberculosis cure rate in developing countries (TB-SMS Cameroon): a protocol of a randomised control study. Trials 2014 Jan 24;15:35 [FREE Full text] [doi: 10.1186/1745-6215-15-35] [Medline: 24460827]
39. Bigna JJR, Kouanfack C, Noubiap JJN, Plottel CS, Koulla-Shiro S. A randomized blinded controlled trial of mobile phone reminders on the follow-up medical care of HIV-exposed and HIV-infected children in Cameroon: study protocol (MORE CARE). Trials 2013 Sep 25;14:313 [FREE Full text] [doi: 10.1186/1745-6215-14-313] [Medline: 24066735]
40. Leon N, Surender R, Bobrow K, Muller J, Farmer A. Improving treatment adherence for blood pressure lowering via mobile phone SMS-messages in South Africa: a qualitative evaluation of the SMS-text Adherence SuppoRt (StAR) trial. BMC Fam Pract 2015 Jul 03;16:80 [FREE Full text] [doi: 10.1186/s12875-015-0289-7] [Medline: 26137844]
41. Hao W, Hsu Y, Chen K, Li H, Iqbal U, Nguyen P, et al. LabPush: a pilot study of providing remote clinics with laboratory results via short message service (SMS) in Swaziland, Africa - a qualitative study. Comput Methods Programs Biomed 2015 Jan;118(1):77-83. [doi: 10.1016/j.cmpb.2014.10.005] [Medline: 25453385]
42. Linnemayr S, Huang H, Luoto J, Kambugu A, Thirumurthy H, Haberer JE, et al. Text messaging for improving antiretroviral therapy adherence: no effects after 1 year in a randomized controlled trial among adolescents and young adults. Am J Public Health 2017 Dec;107(12):1944-1950. [doi: 10.2105/AJPH.2017.304089] [Medline: 29048966]
43. Steury EE. Mobile phone short message service to improve malaria pharmacoadherence in Zambia. J Nurs Scholarsh 2016 Jul;48(4):354-361. [doi: 10.1111/jnu.12216] [Medline: 27145248]
44. Fang R, Deng X. Electronic messaging intervention for management of cardiovascular risk factors in type 2 diabetes mellitus: a randomised controlled trial. J Clin Nurs 2018 Feb;27(3-4):612-620. [doi: 10.1111/jocn.13962] [Medline: 28700102]
45. Zhou H, Sun S, Luo R, Sylvia S, Yue A, Shi Y, et al. Impact of text message reminders on caregivers' adherence to a home fortification program against child anemia in rural western China: a cluster-randomized controlled trial. Am J Public Health 2016 Jul;106(7):1256-1262. [doi: 10.2105/AJPH.2016.303140] [Medline: 27077354]
46. Fang R, Li X. Electronic messaging support service programs improve adherence to lipid-lowering therapy among outpatients with coronary artery disease: an exploratory randomised control study. J Clin Nurs 2016 Mar;25(5-6):664-671. [doi: 10.1111/jocn.12988] [Medline: 26522838]
47. Kumboyono. Short message service as an alternative in the drug consumption evaluation of persons with tuberculosis in Malang, Indonesia. Jpn J Nurs Sci 2017 Apr;14(2):112-116. [doi: 10.1111/jjns.12140] [Medline: 27412502]
48. Kazi AM, Murtaza A, Khoja S, Zaidi AK, Ali SA. Monitoring polio supplementary immunization activities using an automated short text messaging system in Karachi, Pakistan. Bull World Health Organ 2014 Mar 01;92(3):220-225 [FREE Full text] [doi: 10.2471/BLT.13.122564] [Medline: 24700982]
49. Mohan B, Sharma S, Sharma S, Kaushal D, Singh B, Takkar S, et al. Assessment of knowledge about healthy heart habits in urban and rural population of Punjab after SMS campaign-a cross-sectional study. Indian Heart J 2017;69(4):480-484 [FREE Full text] [doi: 10.1016/j.ihj.2017.05.007] [Medline: 28822515]
50. Rico TM, Dos Santos Machado K, Fernandes VP, Madruga SW, Noguez PT, Barcelos CRG, et al. Text messaging (SMS) helping cancer care in patients undergoing chemotherapy treatment: a pilot study. J Med Syst 2017 Oct 09;41(11):181. [doi: 10.1007/s10916-017-0831-3] [Medline: 28990135]
51. Piette JD, Datwani H, Gaudioso S, Foster SM, Westphal J, Perry W, et al. Hypertension management using mobile technology and home blood pressure monitoring: results of a randomized trial in two low/middle-income countries. Telemed J E Health 2012 Oct;18(8):613-620 [FREE Full text] [doi: 10.1089/tmj.2011.0271] [Medline: 23061642]
52. Toda M, Njeru I, Zurovac D, O-Tipo S, Kareko D, Mwau M, et al. Effectiveness of a mobile short-message-service–based disease outbreak alert system in Kenya. Emerg Infect Dis 2016 Apr 1;22(4):711-715 [FREE Full text] [doi: 10.3201/eid2204.151459] [Medline: 26981628]
53. Jia K, Mohamed K. Evaluating the use of cell phone messaging for community Ebola syndromic surveillance in high risked settings in Southern Sierra Leone. Afr Health Sci 2015 Sep;15(3):797-802 [FREE Full text] [doi: 10.4314/ahs.v15i3.13] [Medline: 26957967]
54. Tuijn CJ, Hoefman BJ, van Beijma H, Oskam L, Chevrollier N. Data and image transfer using mobile phones to strengthen microscopy-based diagnostic services in low and middle income country laboratories. PLoS One 2011;6(12):e28348 [FREE Full text] [doi: 10.1371/journal.pone.0028348] [Medline: 22194829]

55. Wu HC, Chang CJ, Lin CC, Tsai MC, Chang CC, Tseng MH. Developing screening services for colorectal cancer on Android smartphones. Telemed J E Health 2014 Aug;20(8):687-695 [FREE Full text] [doi: 10.1089/tmj.2013.0288] [Medline: 24848873]
56. Luchetti M. Global health and the 10/90 gap. Br J Med Pract 2014 Dec 01;7(4):a731 [FREE Full text]

Abbreviations

MeSH: Medical Subject Headings
mHealth: mobile health
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
SMS: short message service

Edited by A Moorhead; submitted 31.12.18; peer-reviewed by S Shah, G Alexander, G Claudine Mbemba, H Seidling; comments to author 10.03.19; revised version received 15.04.19; accepted 25.06.19; published 09.10.19.

Please cite as:
Kruse C, Betancourt J, Ortiz S, Valdes Luna SM, Bamrah IK, Segovia N
Barriers to the Use of Mobile Health in Improving Health Outcomes in Developing Countries: Systematic Review
J Med Internet Res 2019;21(10):e13263
URL: https://www.jmir.org/2019/10/e13263
doi:10.2196/13263
PMID:31593543

Viewpoint

The Role of the Sharing Economy and Artificial Intelligence in Health Care: Opportunities and Challenges

Huailiang Wu[1,2*], MBBS; Nga-Kwo Chan[1,2*], MBBS; Casper J P Zhang[3], MPH, PhD; Wai-Kit Ming[1,2,4], MD, PhD, MPH, MMSc

[1]International School, Jinan University, Guangzhou, China

[2]School of Medicine, Jinan University, Guangzhou, China

[3]School of Public Health, The University of Hong Kong, Hong Kong, China

[4]HSBC Business School, Peking University, Shenzhen, China

[*]these authors contributed equally

Corresponding Author:
Wai-Kit Ming, MD, PhD, MPH, MMSc
School of Medicine
Jinan University
601 West Huangpu Avenue
Guangzhou
China
Phone: 86 14715485116
Email: wkming@connect.hku.hk

Abstract

Health care systems worldwide have been influenced by the globally growing trend toward a sharing economy and will likely advance with these trends in the near future. Therefore, based on peer-to-peer relationships between individuals, sharing health care works by renting medical staff, facilities, and other medical resources. Medical data innovation, integration, analysis, and sharing have the potential to dramatically change the current pattern of the health care system and to provide precise and predictive medical assessment for individuals in the future. In addition, artificial intelligence could be useful in the fields of both clinical medicine and medical research and help to minimize the scarcity of human resources and broaden the role of humans in health care.

(J Med Internet Res 2019;21(10):e13469) doi:10.2196/13469

KEYWORDS

health care; health care system; sharing economy; artificial intelligence

The scarcity of health care resources is a long-standing, persistent global issue that is increasing with the worldwide aging population [1]. Possible approaches toward alleviating this scarcity include applying a sharing economy model to the health care industry [2]. The concept of sharing has been incorporated into a range of commercial activities related to daily life, such as retail and transportation. The health care system has also been influenced by the globally growing trend toward a sharing economy [3] and will likely advance with these trends in the near future. Such foreseeable trends continually accompany the integration of innovative technology in the emerging big-data era, including artificial intelligence (AI). Decision making on global issues requires new technologies based on AI techniques [4]. AI techniques are also needed in clinical prediction, diagnostics, and therapeutics and can be implemented by "learning" from appropriate data (eg, in image-related specialties) [5].

China is currently experiencing rapid economic development and growing health care needs. Given these circumstances, the Chinese health care system seems ideally suited to serve as a testing ground to assess the practicality of advancing health care systems using AI.

Residents of most regions in mainland China are provided with universal health insurance, which means that patients usually only pay a small portion of the treatment fee [6]. In most circumstances, routine treatments involve visiting neighborhood clinics supported by the public sector and then hospitals, if needed. Although reform and development of the primary health care system are proceeding in China, some challenges remain including fragmented health care information and a paucity of

digital data [7]. In Hong Kong, the health care system has a dual-track structure encompassing the public and private sectors [8]. Public health care, operationalized by public hospitals, is the cornerstone of the health care system in Hong Kong, accounting for approximately 90% of health services and 29% of outpatient services. Hong Kong public health care also provides a comprehensive range of quality services at very low costs (eg, US $13/bed/day), which is made possible by a high subsidy rate [7]. In Hong Kong, private clinics complement the public sector. The American system differs from those in mainland China and Hong Kong and involves a hybrid system that includes the federal government, local governments, and private funds (households and private businesses). It is estimated that American citizens have greater health care costs than other developed countries [9]; this is likely because most US health care services are delivered privately, even if they are publicly financed. Selected characteristics (physicians, hospital beds, and mean costs) of the health care systems of mainland China, Hong Kong, and the United States are compared in Table 1. The World Health Organization has specified a minimum of 2.5 health care professionals (physicians, nurses, and midwives) per 1000 people for essential primary care coverage; among the three systems detailed here, only the United States has met this criterion.

Table 1. Characteristics of the selected health care systems (2016).

Characteristic	Mainland China	Hong Kong Special Administrative Region	United States
Physicians per 1000 people	2.31[a]	1.90[b]	2.55[c]
Hospital beds per 1000 people	5.37[a]	5.32[b]	2.5[d]
Mean cost per patient (US $)	1801.74[a]	2208.37[b]	9990[c]

[a]Data source: National Health and Family Planning Commission of the People's Republic of China.

[b]Data source: Department of Health, Hong Kong Special Administrative Region.

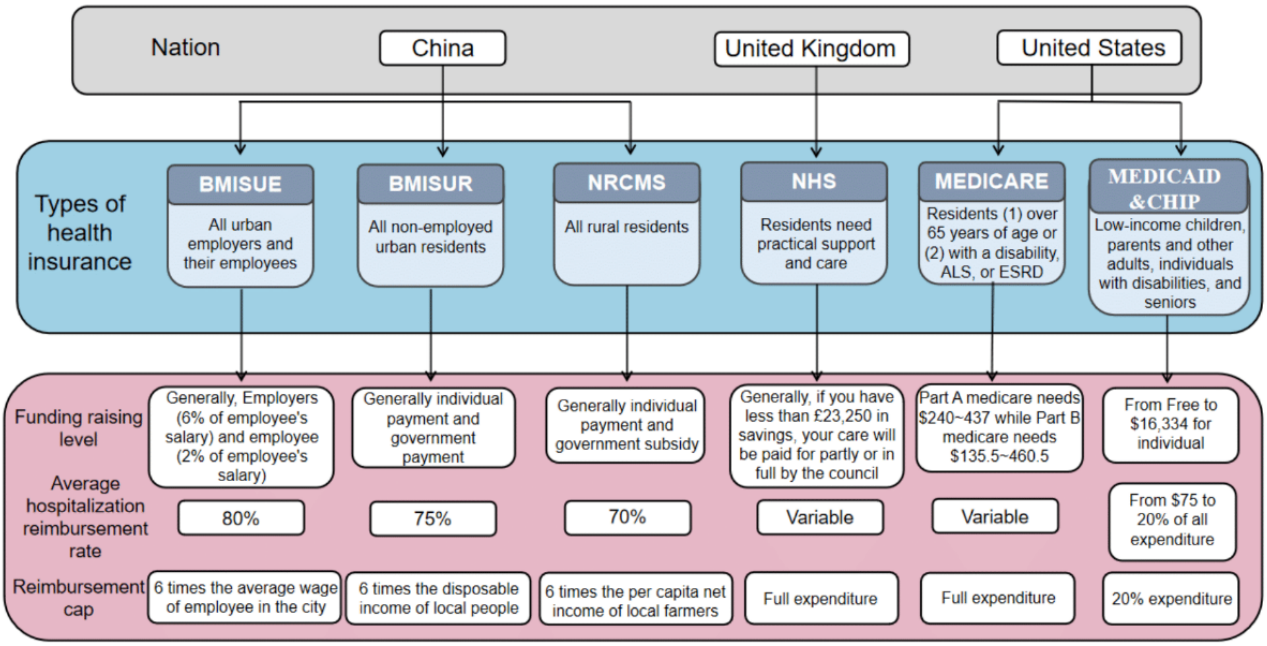
[c]Data source: US Department of Health and Human Services.

[d]Data source: 2014 World Health Organization report.

In a classical health care system, primary care is mainly provided by family physicians. Family physicians are expected to offer appropriate advice and diagnoses at the early stages of illness. This type of health care requires familiarity with patients' conditions (eg, drug allergies or idiopathic issues) and therefore represents a patient-centered approach characterized by preventive and coordinated actions [10]. However, the corresponding costs associated with this comprehensive care tend to be high and can vary across patients. In recent years, the concept of sharing health care has emerged. Sharing health care is based on a peer-to-peer relationship between individuals with similar economic/social roles and works by renting medical staff, facilities, and other medical resources. For example, doctors will practice in single hospitals, clinics, or private health care facilities as well as other locations in a more flexible fashion. This sharing health care system will be similar to other growing sharing businesses (eg, Airbnb, Mobike, and Uber), which are characterized by low entry requirements and high efficiency [2]. In this vision of the future, family physicians will engage in medical record sharing (with informed consent) and paramedic and facility sharing; therefore, with the same limited resources, a larger proportion of patients could receive primary medical care. With the application of principles of sharing economy in medicine, more patients look for medical information on the internet and make decisions based upon that information [11].

The term crowdfunding originally refers to the collective effort of consumers who offer funding to support and sustain a project [12]. Today, an interesting application of sharing economy in health care has emerged in mainland China (eg, among users of Wechat, Weibo, QQ, and other social networking). We consider the application of sharing economy in health care as a type of "charitable crowdfunding," which refers to crowdfunding projects that aim to help the poor, disabled, and other disadvantaged people to pay their health care budgets. As an example, since its foundation in 2015 until September 2018, one popular charitable crowdfunding project in mainland China—"fun in funding"—has already raised more than 55 billion RMB and helped 2.5 million families overcome fatal diseases such as leukemia and lung cancer. Many ordinary people donate just a few RMB, but with the spread of donation messages, more donations follow. As a result, such approaches are likely to raise the targeted money for poor and severely ill patients. Although there is no accurate or reliable statistical data of all funding used in this practice, we believe that this innovation of crowdfunding application has succeeded in helping many low-income families and people with genetic diseases to treat their illnesses. Furthermore, this can help fill the gap of the current health insurance system because all types of health insurance can only cover about 70%–80% of the expenditure if the expenditure does not exceed the reimbursement cap (Figure 1).

Figure 1. A comparison of the Health Insurance Systems of mainland China, the United Kingdom, and the United States. Source for data on BMISUE, BMISUR, and NRCMS is Han and Meng [13]. BMISUE: Basic Medical Insurance System for Urban Employees; BMISUR: Basic Medical Insurance System for Residents; NRCMS: New Rural Cooperative Medical System; NHS: National Health Service; CHIP: Children's Health Insurance Program; ALS: amyotrophic lateral sclerosis; ESRD: end-stage renal disease.



Medical data innovation, integration, analysis, and sharing have the potential to dramatically change the current pattern of the health care system worldwide and provide precise and predictive medical assessment for individuals in the future [14,15]. In addition, decision makers should provide a sufficient medical budget and lead the digitizing health care process through cooperation among the various health care organizations (hospitals, private medical constitutes, and charitable companies) to achieve development of health care systems.

The development of AI has been considerable over the past several decades, and it is conceivable that it could be applied to the medical field, particularly to data evaluation. In about 20 years, half of all work is expected to be out of date or no longer needed, and medical health care is not exempt from AI development [16]. AI could be useful in the fields of clinical medicine and medical research and help minimize the scarcity of human resources and broaden the role of human beings in health care. Given the popularity of personal smart devices, the idea of adopting AI for personal health care services is no longer out of reach and is increasingly viable.

AI technology will indeed help health care professionals provide more care for a larger number of patients, make better clinical decisions, and reduce unnecessary hospitalization and health care costs [17]. However, it is essential to assess the ethical issues in the application of AI technology in health care. Does AI involvement change the doctor-patient relationship and who is responsible for technology-assisted decisions when patients are harmed? Besides, without the clinical physician's involvement, how will patients be reassured and kept emotionally stable? All of these issues require further consideration if effective solutions are to be identified.

When working toward achieving these goals, governments and health care providers must obtain a deeper understanding of patients' preferences of AI in primary care and investigate the applicability of adopting a sharing health care business model.

Conflicts of Interest

None declared.

References

1. Yang Q, Dong H. Have health human resources become more equal between rural and urban areas after the new reform? Int J Health Policy Manag 2014 Dec 03;3(7):359-360 [FREE Full text] [doi: 10.15171/ijhpm.2014.129] [Medline: 25489591]
2. Miller BJ, Moore DW, Schmidt CW. Telemedicine and the sharing economy: the "Uber" for healthcare. Am J Manag Care 2016 Dec 01;22(12):e420-e422 [FREE Full text] [Medline: 27982670]
3. Demiris G. Consumer Health Informatics: Past, Present, and Future of a Rapidly Evolving Domain. Yearb Med Inform 2018 Mar 06;25(S 01):S42-S47. [doi: 10.15265/iys-2016-s005]

4. Fernandez-Luque L, Imran M. Humanitarian health computing using artificial intelligence and social media: A narrative literature review. Int J Med Inform 2018 Jun;114:136-142. [doi: 10.1016/j.ijmedinf.2018.01.015] [Medline: 29395987]
5. Maddox TM, Rumsfeld JS, Payne PRO. Questions for Artificial Intelligence in Health Care. JAMA 2019 Jan 01;321(1):31-32. [doi: 10.1001/jama.2018.18932] [Medline: 30535130]
6. Kong X, Yang Y, Gao J, Guan J, Liu Y, Wang R, et al. Overview of the health care system in Hong Kong and its referential significance to mainland China. J Chin Med Assoc 2015 Oct;78(10):569-573 [FREE Full text] [doi: 10.1016/j.jcma.2015.02.006] [Medline: 26144022]
7. Li X, Lu J, Hu S, Cheng K, De Maeseneer J, Meng Q, et al. The primary health-care system in China. The Lancet 2017 Dec;390(10112):2584-2594. [doi: 10.1016/s0140-6736(17)33109-4]
8. Leung GM, Wong IO, Chan WS, Choi S, Lo SV, Health Care Financing Study Group. The ecology of health care in Hong Kong. Soc Sci Med 2005 Aug;61(3):577-590. [doi: 10.1016/j.socscimed.2004.12.029] [Medline: 15899317]
9. Hood VL, Weinberger SE. High value, cost-conscious care: an international imperative. Eur J Intern Med 2012 Sep;23(6):495-498. [doi: 10.1016/j.ejim.2012.03.006] [Medline: 22863424]
10. Committee on Hospital Care, American Academy of Pediatrics. Family-Centered Care and the Pediatrician's Role. PEDIATRICS 2003 Aug 29;112(3):691-696. [doi: 10.1542/peds.112.3.691]
11. Roper L, Jorm C. Preparing medical students for the e-patient: Is a theoretical grounding required? Medical Teacher 2017 Jul 25;39(10):1101-1101. [doi: 10.1080/0142159x.2017.1353072]
12. Chaney D. A principal–agent perspective on consumer co-production: Crowdfunding and the redefinition of consumer power. Technological Forecasting and Social Change 2019 Apr;141:74-84. [doi: 10.1016/j.techfore.2018.06.013]
13. Han J, Meng Y. Institutional differences and geographical disparity: the impact of medical insurance on the equity of health services utilization by the floating elderly population - evidence from China. Int J Equity Health 2019 Jun 14;18(1):91 [FREE Full text] [doi: 10.1186/s12939-019-0998-y] [Medline: 31200716]
14. Tang C, Plasek JM, Bates DW. Rethinking Data Sharing at the Dawn of a Health Data Economy: A Viewpoint. J Med Internet Res 2018 Nov 22;20(11):e11519 [FREE Full text] [doi: 10.2196/11519] [Medline: 30467103]
15. Moscatelli M, Manconi A, Pessina M, Fellegara G, Rampoldi S, Milanesi L, et al. An infrastructure for precision medicine through analysis of big data. BMC Bioinformatics 2018 Oct 15;19(Suppl 10):351 [FREE Full text] [doi: 10.1186/s12859-018-2300-5] [Medline: 30367571]
16. Frey CB, Osborne MA. The future of employment: How susceptible are jobs to computerisation? Technological Forecasting and Social Change 2017 Jan;114:254-280. [doi: 10.1016/j.techfore.2016.08.019]
17. Meskó B, Hetényi G, Győrffy Z. Will artificial intelligence solve the human resource crisis in healthcare? BMC Health Serv Res 2018 Jul 13;18(1):545 [FREE Full text] [doi: 10.1186/s12913-018-3359-4] [Medline: 30001717]

Abbreviations

AI: artificial intelligence

Edited by G Eysenbach; submitted 23.01.19; peer-reviewed by N Gertrudiz, F Saigí-Rubió; comments to author 27.04.19; revised version received 04.07.19; accepted 19.08.19; published 15.10.19.

Please cite as:
Wu H, Chan NK, Zhang CJP, Ming WK
The Role of the Sharing Economy and Artificial Intelligence in Health Care: Opportunities and Challenges
J Med Internet Res 2019;21(10):e13469
URL: https://www.jmir.org/2019/10/e13469
doi:10.2196/13469
PMID:31617850

Original Paper

Diminishing Effects After Recurrent Use of Self-Guided Internet-Based Interventions in Depression: Randomized Controlled Trial

Lara Bücker[1], MSc; Patricia Schnakenberg[1], MSc; Eirini Karyotaki[2], PhD; Steffen Moritz[1], Prof Dr; Stefan Westermann[1,3], PhD

[1]Department of Psychiatry and Psychotherapy, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

[2]Department of Clinical Psychology, VU Amsterdam, Amsterdam, Netherlands

[3]Idiographic Dynamics Lab, Department of Psychology, University of California, Berkeley, CA, United States

Corresponding Author:
Lara Bücker, MSc
Department of Psychiatry and Psychotherapy
University Medical Center Hamburg-Eppendorf
Martinistr 52
Hamburg, 20246
Germany
Phone: 49 40 7410 55868
Email: l.buecker@uke.de

Abstract

Background: Self-guided internet-based interventions have several advantages over guided interventions and are generally effective in treating psychiatric symptoms.

Objective: We aimed to investigate whether the use of a new self-guided internet-based intervention (MOOD) would lead to a significant reduction in depressive symptoms compared with a care-as-usual (CAU) control group in a sample of individuals with depressive symptoms, most of whom had already used a different self-guided internet-based intervention in a previous trial.

Methods: A total of 125 individuals were randomized to the intervention condition (MOOD) and received access to the intervention for a period of six weeks or a CAU group. After six weeks, all participants were invited to take part in the post assessment. The Beck Depression Inventory-II served as the primary outcome.

Results: Both intention-to-treat as well as per-protocol analyses indicated that the depressive symptomatology decreased in both conditions but showed no advantage for those who had used MOOD. Subsequent moderation analyses suggested that those individuals who had less experience with psychotherapy benefitted to a greater extent compared with those with more experience.

Conclusions: Self-guided internet-based interventions are deemed a suitable first-step approach to the treatment of depression. However, our results indicate that they are more efficacious in those with less psychotherapy experience.

Trial Registration: ClinicalTrials.gov NCT03795480; http://clinicaltrials.gov/ct2/show/NCT03795480

(J Med Internet Res 2019;21(10):e14240) doi:10.2196/14240

KEYWORDS

eHealth; self-management; depressive symptoms; randomized controlled trial

Introduction

Background

Major depression (MD) is one of the most common mental disorders, with more than 300 million people affected worldwide [1]. MD has an aggregated lifetime prevalence of 10.8% and a 1-year prevalence of 7.2% [2], and it represents an enormous personal and economic burden [3,4].

There is evidence for the effectiveness of psychotherapy and pharmacotherapy in the treatment of depression [5-7]. Although classic face-to-face psychotherapy is effective, it is not possible to treat all affected individuals by this method. According to the World Health Organization (WHO), the treatment gap for depression is 56.3% [8]. This means that a large number of

individuals suffering from depressive symptoms remain untreated. This treatment gap can be attributed to several different causes. One of the reasons being that there is an insufficient number of therapists, resulting in long waiting lists. In rural areas, psychologists and psychiatrists are especially underrepresented [9]. Furthermore, general practitioners often do not recognize MD, and misdiagnoses are frequent [10]. Besides these barriers on the supply side, there are also barriers on the demand side. These include patient fear about stigmatization, (expected) discomfort discussing one's own mental health problems, the wish to overcome problems by oneself, a lack of awareness of the need for help, and the anticipated high cost of treatment [11,12]. In addition, it has been shown that greater levels of depression are associated with increased perceived barriers to seeking help, as depression is linked to deficient motivation and reduced activity, as well as to negative cognitive biases (eg, a negative view of the world or future) [13].

The Potential of Internet-Based Interventions

Internet-based interventions can help overcome the supply and demand barriers of conventional face-to-face treatment. Internet-based treatments could reach people who fear stigmatization, are widely accessible (especially for those living in rural areas or those with physical barriers), and provide a high level of privacy [14,15]. Furthermore, the individuals themselves can decide when and where they want to use the intervention. In addition, these interventions are usually available at low cost and, thereby, are affordable for individuals with limited financial resources [16]. In recent years, numerous internet-based interventions have been developed and investigated for their feasibility, acceptance, effectiveness, and side effects, most being used in the treatment of depression and anxiety disorders [17]. To date, a wide range of studies has provided evidence of the effectiveness of internet-based interventions. Studies have found effect sizes ranging from small to large [18-25].

Internet-based interventions can be categorized as either unguided or guided. Guided internet-based interventions are supported by a therapist or a trained person (eg, via frequent email correspondence or telephone support). In self-guided internet-based interventions, the patient does not receive any additional human support. Meta-analyses have found similar effect sizes for guided internet-based treatments when compared with classical face-to-face treatment [18,26]. According to van Ballegooijen et al [27], the average percentage of completed sessions is similar for guided internet-based interventions when compared with face-to-face therapy, although more individuals complete all sessions of the face-to-face therapy. However, individuals who dropped out from face-to-face therapy only completed 24.5% of the intervention, whereas noncompleters of guided internet-based interventions completed on average 42.1%. Many other studies have shown that internet-based interventions often have high dropout rates [28]. A meta-analysis that examined predictors of treatment adherence in self-guided internet-based interventions for depression identified several demographic and psychopathological factors (male gender, low educational background, and comorbid anxiety symptoms) that predict dropout [29].

Another meta-analysis by Richards and Richardson [23] included 19 randomized controlled trials (RCTs) of 10 guided and 9 self-guided interventions for depression. The effect size found for the guided internet-based interventions was 0.78, and for the self-guided interventions the effect size was 0.36 (Cohen *d*). A study that directly compared guided and self-guided interventions for depression within an experimental comparison of the same program with or without therapeutic support and a control group found an effect size of 0.66 for the self-guided version and 1.14 for the guided version (outcome: Beck Depression Inventory-II or BDI-II), which showed that although the effects were smaller for self-guided than for guided, interventions without support were still effective [30]. A meta-analysis by Baumeister et al [31] also supports the superiority of guided over unguided interventions in terms of effectiveness; however, the extent of this superiority seems to be significantly smaller than that in most previous studies (especially in depression studies).

The advantage of self-guided internet-based interventions over guided interventions is that they provide increased access to treatment for those who need it, even for individuals who do not meet the full criteria of a disorder, and, at the same time, are affordable and conserve resources [21]. A recent meta-analysis used individualized participant data to estimate aggregated effect sizes in 13 RCTs on self-guided internet-based interventions for depression [29]. This type of meta-analysis is better able to identify the true effects while taking into account the variability of the studies (eg, degree of support, adherence to treatment, and setting). The meta-analysis found an effect size of 0.27 (Hedge *g*). Contrary to previous results showing that higher baseline symptoms predict a greater reduction of symptoms after the intervention period [32], baseline depressive symptoms did not moderate treatment outcome. Another important finding was that better adherence was associated with better treatment outcome.

Objective

Self-guided internet-based interventions have several advantages over guided interventions and generally are effective in treating psychiatric symptoms. However, the question of which individuals benefit the most has not been investigated well enough. The aim of this study (NCT03795480) was to investigate the acceptance and effectiveness of a new self-guided internet-based intervention for depressive symptoms in a sample that had already received a similar intervention in the context of an earlier study [33]. In other words, we aimed to investigate the possible benefit of recurrent use of self-guided internet-based interventions. The intervention, called MOOD, was developed to provide individuals experiencing subjective depressive symptoms with low-threshold, self-directed, anonymous, and free access to an unguided internet-based intervention. Most of our participants had previously used a Web-based self-help program for depression in the framework of another trial [33]. The study thus enabled us to investigate whether people who had already received a similar therapy could still benefit from MOOD. We expected that participants who received access to MOOD would show a significant reduction of depressive symptoms (primary outcome: BDI-II) compared with the care-as-usual (CAU) control group. In addition, participants in

the intervention group were expected to report a significant increase in self-esteem and quality of life after the intervention period, compared with the control group. Furthermore, willingness to change (as assessed with the University of Rhode Island Change Assessment (URICA) scale [34]) was expected to moderate treatment outcome. Another aim of the study was to examine possible moderators of treatment outcome.

Methods

Study Design

The study was a randomized controlled superiority trial with 2 conditions and parallel assignment (1:1). During the intervention period of 6 weeks, the intervention group received access to the internet-based self-help intervention MOOD, whereas the CAU group received access after completion of the post assessment. There were 2 assessment times, baseline and posttreatment (with an intervention period of 6 weeks). The study was approved by the local psychological ethics committee of the Center for Psychosocial Medicine of the University Medical Clinic Hamburg-Eppendorf, Germany (approval number: LEPEK-003). All participants gave Web-based informed consent before participating in the study. The study was conducted in accordance with the Declaration of Helsinki. The 2 assessments did not ask for any personal information except for an anonymous email address (instructions on how to create such an address were given); no names, telephone numbers, or addresses were asked for. Email addresses were kept in a handwritten list and were assigned to participant codes. The email addresses were stored in a safe. All other obtained data were anonymized and stored electronically on password-protected computers. If a participant requests the deletion of his or her data after completion of the study, this could be done if they provide the code word. The program MOOD ensures network security via secure sockets layer encryption. Messages that were sent within the internal message system of MOOD met the required data safety standard.

Procedure

The study was conducted at the University Medical Center Hamburg-Eppendorf (Germany). At the 2 assessment times, baseline and posttreatment, data were obtained via an internet survey (Enterprise Feedback Suite survey from QuestBack Unipark). The baseline assessment obtained sociodemographic and psychopathological data (see subsection *Instruments*). After 6 weeks, all participants were invited to take part in the post assessment in which the same psychopathological data as before were assessed. Individuals in the treatment group were asked to provide a subjective evaluation of the internet-based self-help intervention. At the beginning of the post assessment, participants were asked to enter the same email address and individual code (numbers and letters) they had generated during the baseline assessment to ensure correct matching of pre and post data. Participants who completed the post assessment were rewarded with access to 2 mindfulness-based self-help manuals. These manuals contain a series of established relaxation and mindfulness exercises (eg, to increase acceptance and self-esteem) and were sent to the participants as PDF files.

Sample size

The power analysis for calculating the sample size for an analysis of covariances (ANCOVA) was conducted using G*Power [35] and revealed a sample size of 128 to detect a medium effect of 0.25, with alpha=.05 and a power of .80, which was expected based on a meta-analysis of the effectiveness of internet-based cognitive behavioral therapy (CBT) for depression [23].

Recruitment

The sample was recruited via an internal database of individuals with depression who had previously participated in studies conducted by the principle investigator's unit and had given consent that they could be contacted for further studies. Most of the participants had previously participated in the Effectiveness of Internet-Based Depression Treatment study (EVIDENT) investigating the effectiveness of an internet-based self-help program for mild to moderate depressive symptoms [33]. Study invitations were sent via email providing information on the purpose and procedure of the study. In addition, invitations to the study were posted on online forums on depression and depression information websites. If participants were interested in taking part in the study, they were directed to a Web page from which the baseline assessment could be started. No financial compensation was offered.

Eligibility Criteria

Participants could be included if they fulfilled the following inclusion criteria: subjective psychological distress with desire for treatment for depressive symptoms (there were no cutoff criteria for depressive symptoms at baseline), aged between 18 and 65 years, internet access, and sufficient command of the German language. Individuals with acute suicidality (assessed at baseline using item 9 on suicidal thoughts of the BDI-II, cutoff ≥2) and/or a self-reported lifetime diagnosis of schizophrenia or bipolar disorder were excluded from the study. Participants who were excluded because of acute suicidality were contacted and provided with help offers and telephone numbers that could be contacted in case of acute crisis. Participants with other psychiatric diagnoses were not excluded from the study. All participants were allowed to continue previously started treatments (psychotherapy or pharmacotherapy; *access to treatment*) and also changes in medication or psychotherapy were allowed during the participation.

Randomization

Participants were randomly allocated into 1 of 2 conditions according to a randomization plan that was set up by the second author using the software Research Randomizer [36]. Block randomization was used to ensure balance between groups. As the study was conducted on the Web and the participants could actively enroll via Web-based registration, the allocation procedure differed from that in classical clinical trials, where allocation is performed by team members. On the basis of the date and time of completion of the baseline assessment, the participants were allocated to conditions following the randomization plan. The allocation rule was 1:1. Participants who were allocated into the intervention group received an email

containing information on the program and a link to the login Web page of MOOD as well as individual login data in the form of a code and a password. Participants in the CAU group received an email with the information that they would receive access to MOOD after completion of the post assessment.

Intervention

During the 6-week intervention period, the intervention group had access to MOOD, an internet-based self-help program targeting depressive symptoms. The program was developed by members of the neuropsychology working group of the University Medical Center Hamburg-Eppendorf and comprised 9 modules (see Table 1). The content of each module is based on CBT techniques and elements of the third wave of CBT.

There is evidence that cognitive restructuring (modules ABC-protocol [A: activating event, B: belief, C: consequence] and modifying thoughts) and behavioral activation (module positive activities) are effective techniques in the treatment of depression [37,38]. In addition, the concept of mindfulness, which has received increasing attention in recent years, is addressed and practiced in a module labeled *mindfulness*. It has been shown that mindfulness has a beneficial effect on the outcome of psychotherapy [39-41]. Strengthening interpersonal skills and competences is strongly recommended within the treatment of depression [37,42] and, therefore, addressed in 1 module (social competence). It is also evident that depression is associated with sleep disturbances, which should, therefore, be targeted in treatment (module sleep) [43,44].

Table 1. Overview of MOOD modules.

Title	Description	Specific skills and exercises
Introduction	Presents the outline and goal of the program; discusses the interactions between thoughts, emotions, and behavior	Introduction into the principles of the program and strengthening of the treatment motivation; creation of list of values and needs according to which the user wants to live; emphasis is placed on the importance of interaction of thoughts, feelings, and behaviors
ABC protocol[a]	Highlights the importance of one's beliefs in dealing with a specific situation; questions automatic beliefs about a situation and helps in developing new, more helpful beliefs	Introduction of the ABC protocol according to Ellis [45]; dysfunctional thought patterns (eg, all-or-nothing thinking and catastrophizing) are explained and converted to more helpful thoughts.
Positive activities	Shows how to integrate positive activities into one's daily routine and achieve one's goals	Presentation of lists of possible positive activities and exercises aimed at integrating positive activities into everyday life on a regular basis and planning them in a meaningful and realistic way; setting short- and long-term goals
Self-esteem	Makes the user aware of his or her own strengths and teaches strategies on how to improve his or her self-perception	The user is asked to identify personal sources of self-esteem and search for forgotten strength. Obstacles are addressed that could stand in the way of an increase of self-esteem (eg, unfair comparisons). Presentation of a list of concrete actions to increase self-esteem in everyday life (eg, joy diary)
Social competence	Presents ways to improve social competences to connect with other individuals and reach goals in social relations	Definition of social competence and presentation of characteristics of aggressive, unsafe, safe, and friendly behavior [46]
Mindfulness	Presents various mindfulness-based relaxation and attention exercises to increase mindfulness in daily life	Identification of signs of mindlessness; distinction between evaluations or judgements and observations to learn an inner attitude of conscious perception, neutrality, and acceptance; suggestions on how the user can distance himself from stressful feelings and obstructive thoughts; presentation of several classical mindfulness exercise (eg, body scan)—those are guided in audio files
Modifying thoughts	Uncovers depressive dysfunctional thoughts and explains methods to turn these into more realistic thoughts	Using ABC protocols to modify depressive dysfunctional thoughts; presentation of different possibilities to positively influence thoughts, such as making concrete statements, avoiding generalizations, changing perspectives, obtaining other opinions, and questioning situations
Sleep	Shows the importance of sleep quality and gives advice on how to improve sleep hygiene	Psychoeducation on the development and maintenance of sleep disorders; identification of characteristics of healthy sleep; creation of a personal list of tips for improving sleep behavior
Relapse prevention	Encourages paying attention to warning signals that might trigger depressive episodes and provides helpful coping strategies	Identification of possible triggers of relapse; presentation of suggestions on how these triggers can be avoided—the user is recommended to balance negative stress with positive activities and identify physical, emotional, cognitive, and behavioral warning signals for stress in oneself (according to a checklist from Kaluza [47]); development of an individual emergency plan

[a]A: activating event, B: belief, C: consequence.

All modules include interactive exercises, worksheets, pictures, graphics, videos, and audios that aim to incorporate the participant's experiences and individual problems to increase the identification of the participant with the material and illustrate the content in an appealing way. The participants were free to choose the order of the modules and could work through the modules at their own speed. We have decided on a free choice of modules to give the participants experiences of autonomy and thus to minimize feelings of heteronomy (eg, being patronized; see [48] for a discussion of the role of motives in internet-based interventions). We recommended that they worked through 1 or 2 modules per week. The approximate time to finish a module ranged between 30 and 60 min. There was no direct guidance. However, the participants had the opportunity to contact a moderator in case of technical questions or problems via messaging within the program. This feature was optional, and the moderator did not actively contact the users on his or her own initiative. However, several reminders were sent via email to those participants who did not log into the program during the study. The participants were encouraged to have a look at the program and work with it. A short overview with summaries of the content of the modules was attached. The emails only served as a reminder to login to the program at least once; no therapeutic support was offered. Of the 55 participants (55/62, 89%) of the MOOD group who logged in to the program, 5 (5/55, 9%) used this feature. A total of 3

participants contacted the moderator several times—2 had 2 contacts (2/55, 4%), 1 had 3 contacts (1/55, 2%). Within the intervention period, none of the participants in the CAU group contacted the study staff.

Instruments

Psychopathological self-rating questionnaires were assessed at the baseline and post assessments. The BDI-II [49] served as the primary outcome. Secondary outcomes included changes in self-esteem and quality of life as well as the participants' subjective evaluation of MOOD.

Beck Depression Inventory-II

The BDI-II [49] was used to assess depressive symptom severity over the previous 2 weeks. The self-rating questionnaire comprises 21 items; for each item, the participant is asked to evaluate the severity of the symptom on a rating scale from 0 to 3, with higher scores indicating more severe depressive symptoms. An overall score of 0 to 13 indicates minimal depression, 14 to 19 indicates mild depression, 20 to 28 indicates moderate depression, and 29 to 63 suggests severe depression. The internal consistency of the BDI-II ranges from 0.79 to 0.90 [50].

Patient Health Questionnaire–9—Depression Module

Change in depressive symptoms was also assessed with the Patient Health Questionnaire–9 (PHQ-9) [51]. The PHQ-9 is a self-rating questionnaire that comprises 9 items on depression, which can be answered on a 4-point rating scale ranging from 0= *not at all* to 3= *nearly every day*. Sum scores can range from 0 to 27 with the following classifications: none or minimal (0-4), mild (5-9), moderate (10-14), and severe (15-27) depressive symptoms. Results of the questionnaire can assist in determining a diagnosis of MD according to Diagnostic and Statistical Manual of Mental Disorders, 4th Edition, criteria. Its internal consistency ranges from 0.86 to 0.89 [50].

Rosenberg Self-Esteem Scale

The Rosenberg Self-Esteem (RSE) scale [52] was used to assess self-esteem. The scale comprises 10 statements regarding self-esteem. Participants are instructed to rate how much they agree with the statements on a 4-point Likert scale from *strongly agree* to *strongly disagree*. Its internal consistency ranges from 0.77 to 0.88. In its original form, higher scores reflect less self-esteem; however, in our study we used a reversed rating scale such that higher scores reflect more self-esteem.

World Health Organization Quality of Life–Abbreviated Version

The WHO Quality of Life–abbreviated version (WHOQOL-BREF) assesses quality of life [53]. It is a short version of the WHOQOL-100, with 26 items. The questionnaire contains 4 different types of 5-point rating scales that ask the participant *how much*, *how complete*, *how often*, *how good*, or *how satisfied* he or she felt over the previous 2 weeks. The questionnaire has 4 subscales: physical health, psychological, social relations, and environment. The WHOQOL-BREF has an internal consistency of 0.70 [54].

University of Rhode Island Change Assessment

The URICA scale is a measure of willingness to change [34] and was used in the baseline survey. It comprises 32 items representing 4 phases of change: precontemplation, contemplation, action, and maintenance. In this study, a total of 9 items were used, selected from the subscales of precontemplation, contemplation, and action. The internal consistency is 0.83, and the reliability of the test-retest lies between 0.63 and 0.75. Answers could be given on a 5-point Likert scale ranging from 1 (strong disagreement) to 5 (strong agreement). In addition, a single item on expectation regarding the treatment outcome was added: "At present, how successful do you think the MOOD self-help program will be?" Here, answers could be given on a 9-point Likert scale ranging from 1 (*not successful at all*) to 9 (*very successful*).

Subjective Appraisal

The subjective appraisal of the intervention was assessed with questions we generated, as well as with adaptions of the Fragebogen zur Patientenzufriedenheit ZUF-8 (questionnaire to measure patient satisfaction [55]). The internal consistency of the ZUF-8 ranges between 0.87 and 0.93. Items could be answered on a 4-point Likert scale ranging from *totally disagree* to *totally agree*. Open questions gave participants the opportunity to provide feedback on the program.

Statistical Analyses

Intention-to-treat (ITT), per-protocol (PP), and frequent user analyses were conducted using IBM Statistics 25 software. The ITT sample comprised all participants who participated in the baseline assessment, whereas the PP sample comprised those who completed both the baseline and the post assessments and used the intervention at least once during the intervention period. Those who used MOOD at least once a week were considered frequent users. According to the Consolidated Standards of Reporting Trials guidelines for reporting RCTs, both ITT and PP analyses should be reported [56]. Although reporting ITT analyses might be considered as the standard analysis for clinical trials, the PP analyses provide an estimate of the true efficacy, as it only includes participants who completed the study and showed (partial) adherence. Missing values in the ITT analysis were imputed using an expectation-maximizing algorithm with 200 imputations. For each measure, we conducted ANCOVA for all samples (ITT, PP, and frequent user) with pre-post differences as the within-group factor, condition as the between-group factor, and baseline scores as the covariate to account for regression toward the mean [57]. In addition, an exploratory moderation analysis was conducted for the PP sample to identify possible moderators that affected differential symptom improvement for the 2 conditions (outcome measure: BDI-II difference scores) using the SPSS macro PROCESS by Hayes [58].

Results

Baseline Characteristics

Table 2 shows the demographic and psychopathological data of the sample at baseline. The first participant was included on May 8, 2018, and the last post assessment took place on July

22, 2018. In total, 125 participants were included in the analyses. Of these, 63 were randomized to the CAU condition and 62 to the MOOD condition (see the study flowchart in Figure 1). Depression symptom severity was on average mild to moderate as measured by the BDI-II: 24/125 (19.2%) had minimal symptom severity (scores 0-13), 29/125 (23.2%) showed mild symptoms (scores 14-19), 39/125 (31.2%) had moderate symptoms (scores 20-28), and 33/125 (26.4%) had severe symptoms (scores 29-63). Of the sample, 49/125 (39.2%) currently were receiving outpatient psychotherapy and 5/125 (4.0%) were waiting for therapy.

Table 2. Demographic, psychopathological, and treatment variables along with respective statistical values (N=125).

Baseline characteristics	MOOD (n=62)	Care-as-usual (n=63)
Demographic characteristics		
Male, n (%)	15 (24)	18 (29)
Age (years), mean (SD)	44.02 (10.90)	48.02 (10.95)
Level of school education (% A-level), n (%)	37 (60)	36 (57)
Treatment variables		
Length of distress (years), mean (SD)	11.63 (7.97)	11.47 (7.91)
Medication, n (%)		
None	38 (61)	35 (56)
Antidepressant	20 (32)	22 (35)
Antidepressant and antipsychotics	4 (7)	6 (10)
Treatment status, n (%)		
None	27 (44)	28 (44)
In treatment	24 (39)	25 (40)
Waiting for therapy	2 (3)	3 (5)
Treatment expectation[a], mean (SD)	5.26 (1.55)	5.63 (1.42)
Number of previous courses of psychotherapy, n (%)		
0	4 (7)	9 (14)
1-2	34 (55)	23 (37)
>2	24 (39)	31 (49)
Psychometric scales, mean (SD)		
BDI-II[b] (depressive symptom severity)	22.54 (11.39)	22.79 (11.98)
PHQ-9[c] (depressive symptom severity)	10.74 (4.50)	10.52 (5.17)
WHOQOL-BREF[d]		
Global	48.39 (21.88)	49.80 (21.47)
Physical health	55.59 (18.27)	57.31 (19.18)
Psychological	43.55 (18.74)	45.90 (18.48)
Social relationships	45.43 (20.56)	48.41 (20.62)
Environmental	67.69 (16.47)	73.61 (16.01)
RSE[e] (self-esteem)	26.14 (6.96)	26.92 (6.92)

[a]1=*not at all successful* to 9=*very successful*.

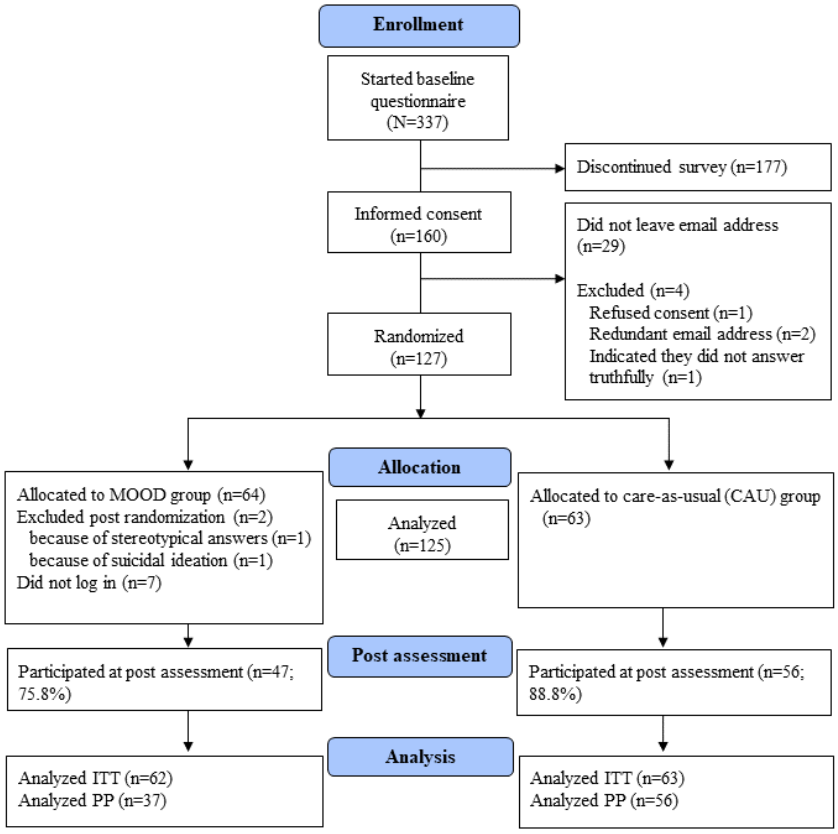
[b]BDI-II: Beck Depression Inventory-II.

[c]PHQ-9: Patient Health Questionnaire–9.

[d]WHOQOL-BREF: World Health Organization Quality of Life–abbreviated version.

[e]RSE: Rosenberg Self-Esteem.

Figure 1. Consolidated Standards of Reporting Trials flow diagram. ITT: intention-to-treat; PP: per-protocol.



Completion of Assessments and Treatment Adherence

The completion rate of the assessments was satisfactory; 103/125 (82.4%) completed the post assessment. There were no significant differences in completion of assessments between the intervention group and the CAU group ($\chi^2_{1,125}$=3.7; *P*=.06). Completers and noncompleters were not significantly different on any baseline demographic characteristic except for 2 items regarding the presence of a diagnosis of a depressive disorder and substance or alcohol dependency. Specifically, of those who did not complete the post assessment, 6/22 (27%) did not have a self-reported diagnosis of depression, whereas of those who completed the post assessment, 9/103 (8.7%) did not have a self-reported diagnosis of depression ($\chi^2_{1,125}$=5.9; *P*=.02). Relatively more participants, 3/22 (14%), who did not complete the post assessment indicated having a diagnosis of a substance or alcohol dependency compared with those who completed the study and had such a diagnosis (3/103, 2.9%; $\chi^2_{1,125}$=4.6; *P*=.03).

Although the completion rate of the assessments was satisfactory, the usage of the intervention was rather low. Only 24/62 (39%) logged into the program at least once a week (frequent users). Of those who logged into the program, the participants completed an average of 2.90 (SD 3.13) modules. The mean time (in minutes) the users engaged with the program was 117.36 min (SD 209.05). No significant correlation could be found between duration of use and the number of completed modules with symptom improvement after the intervention (*P*>.05).

Primary Outcome

Tables 3 and 4 shows the group differences across time of each measure for the PP sample, frequent user sample (those who logged into the program at least once a week), and ITT sample. For the primary outcome (score on the BDI-II), the results of the ANCOVA analysis did not reach statistical significance, neither for ITT nor PP or frequent user sample. Results of the paired sample *t* tests showed a significant decline of depressive symptoms from pre to post for the intervention (t_{46}=2.05; *P*=.05) and CAU (t_{55}=2.96; *P*=.004) groups.

Table 3. Group differences across time; means, standard deviations, effect sizes (Cohen *d*) and 95% CIs of completer sample (within-group differences are denoted via superscripts).

Measurements	MOOD			Care-as-usual		
	Pre (n=62), mean (SD)	Post (n=47), mean (SD)	Cohen *d* (95% CI)	Pre (n=63), mean (SD)	Post (n=56), mean (SD)	Cohen *d* (95% CI)
BDI-II[a]	22.54 (11.39)	20.36 (14.70)	−0.17 (−0.55 to 0.21)[b]	22.79 (11.99)	18.68 (12.79)	−0.35 (−0.71 to 0.02)[c]
PHQ-9[d]	10.74 (4.50)	10.60 (6.23)	−0.03 (−0.41 to 0.35)	10.52 (5.17)	9.11 (5.64)	−0.26 (−0.62 to 0.01)[b]
WHOQOL-BREF[e]						
Global	48.39 (21.88)	50.00 (19.85)	0.08 (−0.30 to 0.46)	49.80 (21.47)	55.13 (22.71)	0.24 (−0.12 to 0.60)[b]
Physical health	55.59 (18.28)	57.22 (20.33)	−0.09 (−0.29 to 0.46)	57.31 (19.18)	60.33 (19.89)	0.16 (−0.21 to 0.52)
Psychological	43.55 (18.74)	46.72 (20.43)	−0.16 (−0.22 to 0.54)[f]	45.90 (18.48)	49.70 (19.54)	0.20 (−0.16 to 0.56)[b]
Social relationships	45.43 (20.56)	48.40 (23.03)	−0.14 (−0.24 to 0.52)	48.41 (20.62)	50.30 (20.59)	0.09 (−0.27 to 0.45)
Environment	67.69 (16.47)	70.01 (14.09)	−0.15 (−0.23 to 0.53)[b]	73.61 (16.01)	76.28 (15.17)	0.17 (−0.19 to 0.53)[b]
RSE[g]	26.15 (6.96)	27.34 (7.67)	−0.16 (−0.22 to 0.54)[c]	26.92 (6.12)	28.68 (6.74)	0.27 (−0.09 to 0.64)[c]

[a]BDI-II: Beck Depression Inventory-II.

[b]$P \leq .05$.

[c]$P \leq .005$.

[d]PHQ-9: Patient Health Questionnaire–9.

[e]WHOQOL-BREF: World Health Organization Quality of Life–abbreviated version.

[f]$P \leq .01$

[g]RSE: Rosenberg Self-Esteem.

Table 4. Analysis of covariances with respective baseline values as covariates.

Measurements	Per protocol sample (n=93)		*P* value	Frequent user sample (n=80)		*P* value	Intention to treat (n=125)		*P* value
	$F_{1,90}$	η_p^2		$F_{1,77}$	η_p^2		$F_{1,122}$	η_p^2	
BDI-II[a]	0.24	0.003	.63	0.24	0.003	.63	0.94	0.008	.34
PHQ-9[b]	0.03	0.000	.86	0.00	0.000	.97	2.41	0.019	.12
WHOQOL-BREF[c]									
Global	0.10	0.001	.75	0.03	0.000	.87	1.72	0.014	.19
Physical health	0.66	0.007	.42	0.04	0.001	.84	0.37	0.003	.55
Psychological	0.57	0.006	.45	1.75	0.022	.19	0.11	0.001	.75
Social relationships	0.69	0.008	.41	0.33	0.004	.57	0.10	0.001	.75
Environment	0.01	0.000	.95	0.27	0.003	.61	0.92	0.008	.34
RSE[d]	0.879	0.010	.35	0.51	0.007	.48	0.44	0.004	.51

[a]BDI-II: Beck Depression Inventory-II.

[b]PHQ-9: Patient Health Questionnaire–9.

[c]WHOQOL-BREF: World Health Organization Quality of Life–abbreviated version.

[d]RSE: Rosenberg Self-Esteem.

Secondary Outcomes

Results of paired sample *t* tests indicate a significant increase in scores on the self-esteem scale (RSE) for the intervention (t_{46}=2.97; *P*=.005) and CAU (t_{55}=3.04; *P*=.004) groups. There was no significant improvement across time in scores on the RSE as analyzed using an ANCOVA with baseline score as covariate (see Table 4). This is applicable to all samples analyzed (ITT, PP, and frequent user). Quality of life was assessed with the WHOQOL-BREF. Although there was a significant increase in the global score of the CAU control group (t_{55}=2.46; *P*=.02), there was no significant increase in the intervention group (t_{46}=1.09; *P*=.28). Paired sample *t* tests for the subscales of the WHOQOL-BREF (physical health, psychological, social relationships, and environment) found a significant increase on the WHOQOL-BREF psychological

(MOOD: t_{46}=−2.86; *P*=.006 and CAU: t_{55}=−2.50; *P*=.02) and environmental (MOOD: t_{46}=−2.08; *P*=.04 and CAU: t_{55}=−2.27; *P*=.03) scales for both groups. The ANCOVA analysis did not find a significant difference across groups for the WHOQOL-BREF and its subscales.

Moderation Analysis

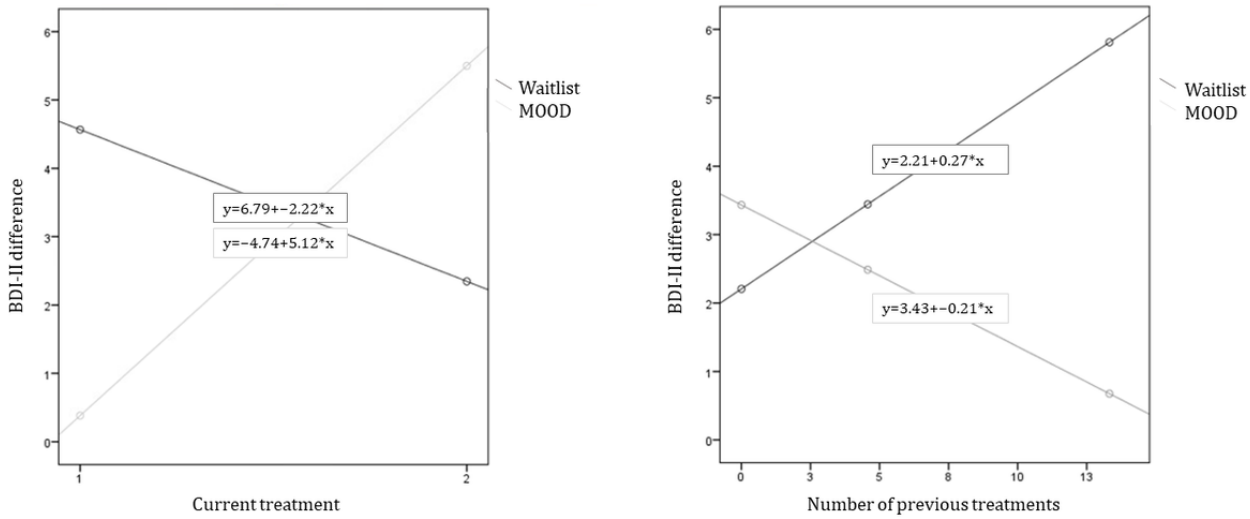
The results for the interaction effect of the exploratory moderation analyses are summarized in Table 5. Subsequent analyses showed that participants in the intervention group who had less experience with psychotherapy showed a greater improved outcome on depression (BDI-II) compared with the CAU group (*P*=.03). In addition, moderation analyses showed results that bordered on significance (*P*=.05), suggesting that individuals in the intervention group who were currently receiving no other treatment had a better outcome compared with the CAU group (see Figure 2 and Table 5). Results indicate that willingness to change did not moderate treatment outcome (*P*>.05).

Table 5. Moderators for improvement in depression (Beck Depression Inventory-II difference score, means are centered).

Outcome parameter	Beta	SE	*t*	*P* value	95% CI
Number of past courses of psychotherapy	−.478	0.219	−2.178	.03	−0.913 to −0.042
Current treatment[a]	7.340	3.711	1.978	.05	−0.035 to 14.714

[a]Possible answers were no therapy, in-patient treatment, out-patient treatment, treatment with licensed psychotherapist, treatment at day hospital, waiting for therapy to begin, and planning on starting therapy.

Figure 2. Interaction effects of current treatment (left), number of prior treatment (right) and group allocation. The graph on the left presents the effects of current treatment (1=yes, 2=no) on symptom reduction. The graph on the right depicts how experience with psychotherapy (number of prior treatments) is related to depressive symptomatology (outcome: reduction on BDI-II). BDI-II: Beck Depression Inventory-II.



Subjective Appraisal

In Tables 6 and 7, subjective appraisals of MOOD are displayed. Overall, MOOD was positively evaluated. Of those who used the intervention, 78% (29/37) rated MOOD as suitable for self-help, 73% (27/37) considered it an applicable supplement to psychotherapy, and 54% (20/37) rated the program as helpful. Furthermore, 81% (30/37) rated the quality of the program as excellent to good, 76% (28/37) would recommend it to a friend with similar problems, and 68% (25/37) would use the program again. However, only 22% (8/37) indicated that they were able to use MOOD regularly, 46% (17/37) had to force themselves to use the program, and only 30% (11/37) stated that their depressive symptoms decreased through using MOOD. In addition, 60% (18/30) rated MOOD to be equally good or better than the internet-based intervention they had previously used, and 17% (5/30) considered MOOD inferior.

Table 6. Subjective appraisal of MOOD (scores: 1=not at all, 2=a little, 3=a lot, and 4=absolutely).

Item	MOOD condition (n=37), mean (SD)	Positive (*a lot* and *absolutely*) appraisal, n (%)
I think the MOOD program is good for self-help and self-guidance.	2.95 (0.85)	29 (78)
I think the contents of the program were understandable.	1.97 (0.93)	35 (95)
I think the program was helpful.	2.54 (1.02)	20 (54)
I was able to use the program on a regular basis during the past 6 weeks.	1.86 (0.95)	8 (22)
I had to force myself to use the program.	2.46 (1.17)	17 (46)
My depressive symptoms decreased because of the use of the program.	1.97 (0.93)	11 (30)
I consider the program to be applicable as a supplement to psychotherapy.	2.97 (0.87)	27 (73)
The program is not applicable to my depressive symptoms.	1.76 (1.07)	9 (24)

Table 7. Subjective appraisal of MOOD (adapted from a German questionnaire on patient satisfaction, ZUF-8 [59]).

Item	MOOD condition (n=37), mean (SD)	Positive appraisal, n (%)
How do you rate the quality of the program? (excellent, good vs okay, not good)	2.30 (1.05)	30 (81)
Did you receive the type of treatment you expected to receive? (absolutely, a lot vs a little, not at all)	3.16 (0.90)	25 (68)
To what extent did the program help you cope with your problems? (absolutely, a lot vs a little, not at all)	2.76 (1.14)	19 (51)
Would you recommend the program to a friend with similar symptoms? (yes, probably yes vs probably not, no)	3.32 (0.97)	28 (76)
How happy are you about the extent of the help you have received through using the program? (very satisfied, mostly satisfied vs somewhat dissatisfied, dissatisfied)	3.16 (1.07)	21 (57)
Did the program help you cope with your problems more successfully? (absolutely, a lot vs a little, not at all)	2.65 (1.36)	21 (57)
How satisfied are you with the program in general? (very satisfied, mostly satisfied vs somewhat unsatisfied, unsatisfied)	2.57 (1.28)	25 (68)
Would you use the program again? (Yes, probably yes vs probably not, no)	3.35 (1.11)	25 (68)

Discussion

Principal Findings

The treatment gap in the treatment of depressive disorders is a significant issue [8] and might be narrowed by providing effective internet-based interventions. Self-guided internet-based interventions can be especially useful because therapists are a scarce resource. Furthermore, depressive episodes tend to reoccur [60], but treatment barriers are stable, which means that affected individuals are repeatedly confronted with the problem of the undersupply of therapists. In our study, we aimed to investigate whether the use of a new self-guided internet-based intervention called MOOD over a period of 6 weeks would lead to a significant reduction in depressive symptoms compared with a CAU group (that got access to MOOD after completion of the post assessment) in a sample of individuals who had previously and/or currently received therapy (most of the participants had also already received a different internet-based intervention).

In our analyses, we found no significant difference across groups in any outcome variable (for ITT, PP, and frequent users). Contrary to our hypothesis, the group that received MOOD over the intervention period did not significantly improve in depressive symptoms compared with the CAU group. Both groups significantly improved over time and, interestingly, the difference here was more pronounced for the CAU group for the primary outcome, which was the BDI-II. Both groups significantly improved over time in levels of self-esteem and quality of life, but this effect again was not different across groups. These findings are different from previous trials investigating the effectiveness of self-guided internet-based interventions. Here, it was found that self-guided internet-based interventions were effective in treating depressive symptoms at a small to medium effect size [21,23]. Certainly, it is possible that this difference between ours and previous results might stem from a publication bias. Furthermore, self-guided internet-based interventions were found to be more effective in individuals currently not seeing a psychotherapist [33,61]. Results of this study support this finding with a P value of .05, which indicates that individuals in the intervention group who benefited more from the program were currently not receiving other psychotherapy. However, this result was not statistically significant. The number of participants who were in treatment while using MOOD amounted to 39%, and these individuals had a smaller symptom reduction compared with those who

were not in treatment. This result indicates that internet-based interventions appear to be more useful to individuals who are currently not receiving psychotherapy. More than 70% of participants in the intervention group considered MOOD to be a suitable supplement to psychotherapy. However, this is more likely to be the case if the combination of internet-based interventions and psychotherapy is addressed in therapy, and the integration is actively promoted, as in blended therapy [62,63,64].

Another interesting finding was that the participants' past experience with psychotherapy had a significant effect on treatment outcome. Here, results indicated that individuals with less experience benefited more than those with more experience. Our sample mainly comprised individuals with a relatively long history of depression; most of them had undergone multiple courses of psychotherapy in the past. Only 10.4% of the overall sample had not previously received psychotherapy. Also, 82.9% had previously used a comparable self-guided internet-based intervention within the framework of another trial [33]. Our results support the assumption that individuals who have already used a very similar internet-based intervention and have a high level of psychotherapeutic experience may not benefit from another self-guided Web-based intervention. Perhaps they should receive personalized support, such as a guided internet-based intervention. It is conceivable that individuals who have suffered from depression for a long time and who already have had significant treatment experience will not learn much that is unknown or new to them in a subsequent intervention. Another explanation might be that individuals who have not benefited from conventional psychotherapy and, therefore, have continued to look for other forms of treatment also will not benefit from internet-based treatment. The fact that people with a long history of depression usually suffer from residual symptoms that are difficult to treat is a well-known prognostic factor and therefore is presumably another reason for the nonsignificant group differences we found in our sample [65,66]. Another important point is that patient's expectations of the outcome of therapy is considered a key determinant of the actual therapy outcome [67,68]. Many patients with a long history of a mental disorder and extensive therapy experience are likely to have lower expectations and hopes regarding treatment outcome. This naturally applies to other forms of therapy too and is not confined to internet-based interventions. However, it could be a further reason for the lack of treatment effects found in our study. Despite this, based on the results presented above, we come to the conclusion that unguided internet-based interventions should mainly serve as a first step in treatment and are not suitable for individuals with recurrent or residual symptoms. Our findings, therefore, support stepped care approaches, in which self-guided internet-based interventions are used in early phases of illness and in patients without a chronification of symptoms [69-71]. The key idea behind these stepped models is that patients should first be treated at the lowest level (to avoid *psychiatrization* and self-stigma) and only proceed to more advanced care when symptoms are severe or do not improve with (guided) self-help. It has been found that stepped care approaches can be effective and resource-saving in the treatment of common mental disorders, such as depression [72]. Another potential implication is that programs for individuals with therapeutic experience have to be more elaborated to be effective. For instance, they could consider motives of participants [48] or be personalized in other ways.

Strengths and Limitations

To participate in the study, neither a (verified) diagnosis of depression nor a minimum severity of depression symptoms was necessary, which resulted in heterogeneity of depression levels. Therefore, on the one hand, a wider range of individuals with desire for treatment for depression was reached, regardless of whether they met the criteria of a diagnosis or not. On the other hand, it has been found that samples of severely depressed participants benefit more from low-intensity psychological interventions than samples of mildly depressed participants [32], which might be because those with severe depression have more room to improve. This finding, however, contrasts with results of Karyotaki et al [21] who found that self-guided internet-based interventions are effective regardless of symptom severity. Interestingly, within another study by Karyotaki et al [73], it was found that individuals with more severe baseline symptoms were more likely to improve than individuals with less severe baseline symptoms after treatment with guided internet-based interventions, which means that the findings are still ambiguous in this respect. Apart from this, broad inclusion criteria may result in a type I error and, thereby, lead to an underestimation of the intervention's treatment potential. Another limitation might be the sample size in our study. Our power analysis was based on a medium effect, and we did not consider expected dropout in our sample size calculation, which could have led to an overestimation, as Karyotaki et al [21] only found a small effect (g=0.27) for self-guided interventions. To detect such a small effect, a larger sample would have been necessary.

A strength of the study is the high completion rate of the assessments—only 17.6% did not participate or complete the post assessment. Despite the high completion rate, the treatment adherence was rather low; only 39% were able to use the program on a regular basis (at least once a week), and the average time of program usage was 117.36 min. Contrary to the results of other studies, there was no significant correlation between duration of use and therapy outcome [21,74]. Interestingly, the percentage of participants who self-reported a regular use was minor (22%) compared with the percentage of the frequent users as defined by us. It is conceivable that depressive users themselves define much higher standards for regular use. If they do not meet these demands, this leads to disappointment and negative feelings. This point can be discussed in the context of depression-specific negative cognitive distortions, namely, (dysfunctional) perfectionism and *all-or-nothing thinking* [75,76]. Individuals with depression, who have these thought distortions, often think that an expectation task is either fulfilled perfectly (100%) or not at all. In our case, this could mean that the participants rated themselves or their use of the program worse than they actually did because of their depressive way of thinking. Strategies that improve adherence in self-guided internet-based interventions (such as sending reminders) have been found to have a positive effect on treatment outcome [77]. However, in our study, we did send frequent reminders to participants to use the

intervention. Finally, as the study had limited financial resources, follow-up assessments were not possible. For this reason, no conclusion can be drawn about long-term effects.

Conclusions

The efficacy of unguided internet-based interventions has been proven in various trials. However, it is unclear whether internet-based interventions are beneficial for individuals with a long history of depression and greater experience with psychotherapy and for those who have already undergone a different unguided Web-based intervention. With this sample, we could not verify the efficacy of the new self-help internet-based intervention MOOD; however, 2 interesting findings with implications for the future were identified. Greater experience with psychotherapy and (at trend level) current treatment reduced the effects of the intervention on depressive symptoms. We conclude that unguided internet-based interventions might be appropriate as a first step in treatment but not for individuals who already have much experience with psychotherapy or internet-based interventions.

Acknowledgments

The study was partly supported by a grant of the Swiss National Science Foundation to SW and SM (Grant Number 159384) and a fellowship of the Swiss National Science Foundation to SW (Number 177678). The study was not externally funded. It was financially supported by private donations to the working group and by overhead funds of the institution.

Conflicts of Interest

None declared.

Multimedia Appendix 1
CONSORT - EHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File)107 KB - jmir_v21i10e14240_app1.pdf]

References

1. World Health Organization. 2017. Depression and Other Common Mental Disorders Global Health Estimates URL: https://apps.who.int/iris/bitstream/handle/10665/254610/WHO-MSD-MER-2017.2-eng.pdf [accessed 2019-08-15]
2. Lim GY, Tam WW, Lu Y, Ho CS, Zhang MW, Ho RC. Prevalence of depression in the community from 30 countries between 1994 and 2014. Sci Rep 2018 Feb 12;8(1):2861 [FREE Full text] [doi: 10.1038/s41598-018-21243-x] [Medline: 29434331]
3. Sobocki P, Jönsson B, Angst J, Rehnberg C. Cost of depression in Europe. J Ment Health Policy Econ 2006 Jun;9(2):87-98. [Medline: 17007486]
4. Lépine JP, Briley M. The increasing burden of depression. Neuropsychiatr Dis Treat 2011;7(Suppl 1):3-7 [FREE Full text] [doi: 10.2147/NDT.S19617] [Medline: 21750622]
5. Cuijpers P, Berking M, Andersson G, Quigley L, Kleiboer A, Dobson KS. A meta-analysis of cognitive-behavioural therapy for adult depression, alone and in comparison with other treatments. Can J Psychiatry 2013 Jul;58(7):376-385. [doi: 10.1177/070674371305800702] [Medline: 23870719]
6. Cuijpers P, Andersson G, Donker T, van Straten A. Psychological treatment of depression: results of a series of meta-analyses. Nord J Psychiatry 2011 Dec;65(6):354-364. [doi: 10.3109/08039488.2011.596570] [Medline: 21770842]
7. Cuijpers P, Sijbrandij M, Koole SL, Andersson G, Beekman AT, Reynolds CF. The efficacy of psychotherapy and pharmacotherapy in treating depressive and anxiety disorders: a meta-analysis of direct comparisons. World Psychiatry 2013 Jun;12(2):137-148 [FREE Full text] [doi: 10.1002/wps.20038] [Medline: 23737423]
8. Kohn R, Saxena S, Levav I, Saraceno B. The treatment gap in mental health care. Bull World Health Organ 2004 Nov;82(11):858-866 [FREE Full text] [doi: 10.1590/S0042-96862004001100011] [Medline: 15640922]
9. Andrade LH, Alonso J, Mneimneh Z, Wells JE, Al-Hamzawi A, Borges G, et al. Barriers to mental health treatment: results from the WHO World Mental Health surveys. Psychol Med 2014 Apr;44(6):1303-1317 [FREE Full text] [doi: 10.1017/S0033291713001943] [Medline: 23931656]
10. Carey M, Jones K, Meadows G, Sanson-Fisher R, D'Este C, Inder K, et al. Accuracy of general practitioner unassisted detection of depression. Aust N Z J Psychiatry 2014 Jun;48(6):571-578 [FREE Full text] [doi: 10.1177/0004867413520047] [Medline: 24413807]
11. Clement S, Schauman O, Graham T, Maggioni F, Evans-Lacko S, Bezborodovs N, et al. What is the impact of mental health-related stigma on help-seeking? A systematic review of quantitative and qualitative studies. Psychol Med 2015 Jan;45(1):11-27. [doi: 10.1017/S0033291714000129] [Medline: 24569086]
12. Gulliver A, Griffiths KM, Christensen H. Perceived barriers and facilitators to mental health help-seeking in young people: a systematic review. BMC Psychiatry 2010 Dec 30;10:113 [FREE Full text] [doi: 10.1186/1471-244X-10-113] [Medline: 21192795]

13. Mohr DC, Ho J, Duffecy J, Baron KG, Lehman KA, Jin L, et al. Perceived barriers to psychological treatments and their relationship to depression. J Clin Psychol 2010 Apr;66(4):394-409 [FREE Full text] [doi: 10.1002/jclp.20659] [Medline: 20127795]
14. Lal S, Adair CE. E-mental health: a rapid review of the literature. Psychiatr Serv 2014 Jan 1;65(1):24-32. [doi: 10.1176/appi.ps.201300009] [Medline: 24081188]
15. Richards D, Viganó N. Online counseling: a narrative and critical review of the literature. J Clin Psychol 2013 Sep;69(9):994-1011. [doi: 10.1002/jclp.21974] [Medline: 23630010]
16. Andersson G, Titov N. Advantages and limitations of internet-based interventions for common mental disorders. World Psychiatry 2014 Feb;13(1):4-11 [FREE Full text] [doi: 10.1002/wps.20083] [Medline: 24497236]
17. Hedman E, Ljótsson B, Lindefors N. Cognitive behavior therapy via the internet: a systematic review of applications, clinical efficacy and cost-effectiveness. Expert Rev Pharmacoecon Outcomes Res 2012 Dec;12(6):745-764. [doi: 10.1586/erp.12.67] [Medline: 23252357]
18. Carlbring P, Andersson G, Cuijpers P, Riper H, Hedman-Lagerlöf E. Internet-based vs face-to-face cognitive behavior therapy for psychiatric and somatic disorders: an updated systematic review and meta-analysis. Cogn Behav Ther 2018 Jan;47(1):1-18. [doi: 10.1080/16506073.2017.1401115] [Medline: 29215315]
19. Ebert DD, Donkin L, Andersson G, Andrews G, Berger T, Carlbring P, et al. Does internet-based guided-self-help for depression cause harm? An individual participant data meta-analysis on deterioration rates and its moderators in randomized controlled trials. Psychol Med 2016 Oct;46(13):2679-2693 [FREE Full text] [doi: 10.1017/S0033291716001562] [Medline: 27649340]
20. Karyotaki E, Ebert DD, Donkin L, Riper H, Twisk J, Burger S, et al. Do guided internet-based interventions result in clinically relevant changes for patients with depression? An individual participant data meta-analysis. Clin Psychol Rev 2018 Jul;63:80-92. [doi: 10.1016/j.cpr.2018.06.007] [Medline: 29940401]
21. Karyotaki E, Riper H, Twisk J, Hoogendoorn A, Kleiboer A, Mira A, et al. Efficacy of self-guided internet-based cognitive behavioral therapy in the treatment of depressive symptoms: a meta-analysis of individual participant data. JAMA Psychiatry 2017 Apr 1;74(4):351-359. [doi: 10.1001/jamapsychiatry.2017.0044] [Medline: 28241179]
22. Lüdtke T, Westermann S, Pult LK, Schneider BC, Pfuhl G, Moritz S. Evaluation of a brief unguided psychological online intervention for depression: a controlled trial including exploratory moderator analyses. Internet Interv 2018 Sep;13:73-81 [FREE Full text] [doi: 10.1016/j.invent.2018.06.004] [Medline: 30206522]
23. Richards D, Richardson T. Computer-based psychological treatments for depression: a systematic review and meta-analysis. Clin Psychol Rev 2012 Jun;32(4):329-342. [doi: 10.1016/j.cpr.2012.02.004] [Medline: 22466510]
24. Schröder J, Berger T, Westermann S, Klein JP, Moritz S. Internet interventions for depression: new developments. Dialogues Clin Neurosci 2016 Jun;18(2):203-212 [FREE Full text] [Medline: 27489460]
25. Stolz T, Schulz A, Krieger T, Vincent A, Urech A, Moser C, et al. A mobile app for social anxiety disorder: a three-arm randomized controlled trial comparing mobile and PC-based guided self-help interventions. J Consult Clin Psychol 2018 Jun;86(6):493-504. [doi: 10.1037/ccp0000301] [Medline: 29781648]
26. Andersson G, Cuijpers P, Carlbring P, Riper H, Hedman E. Guided internet-based vs face-to-face cognitive behavior therapy for psychiatric and somatic disorders: a systematic review and meta-analysis. World Psychiatry 2014 Oct;13(3):288-295 [FREE Full text] [doi: 10.1002/wps.20151] [Medline: 25273302]
27. van Ballegooijen W, Cuijpers P, van Straten A, Karyotaki E, Andersson G, Smit JH, et al. Adherence to internet-based and face-to-face cognitive behavioural therapy for depression: a meta-analysis. PLoS One 2014;9(7):e100674 [FREE Full text] [doi: 10.1371/journal.pone.0100674] [Medline: 25029507]
28. Melville KM, Casey LM, Kavanagh DJ. Dropout from internet-based treatment for psychological disorders. Br J Clin Psychol 2010 Nov;49(Pt 4):455-471. [doi: 10.1348/014466509X472138] [Medline: 19799804]
29. Karyotaki E, Kleiboer A, Smit F, Turner DT, Pastor AM, Andersson G, et al. Predictors of treatment dropout in self-guided web-based interventions for depression: an 'individual patient data' meta-analysis. Psychol Med 2015 Oct;45(13):2717-2726. [doi: 10.1017/S0033291715000665] [Medline: 25881626]
30. Berger T, Hämmerli K, Gubser N, Andersson G, Caspar F. Internet-based treatment of depression: a randomized controlled trial comparing guided with unguided self-help. Cogn Behav Ther 2011;40(4):251-266. [doi: 10.1080/16506073.2011.616531] [Medline: 22060248]
31. Baumeister H, Reichler L, Munzinger M, Lin J. The impact of guidance on internet-based mental health interventions — a systematic review. Internet Interv 2014 Oct;1(4):205-215. [doi: 10.1016/j.invent.2014.08.003]
32. Bower P, Kontopantelis E, Sutton A, Kendrick T, Richards DA, Gilbody S, et al. Influence of initial severity of depression on effectiveness of low intensity interventions: meta-analysis of individual patient data. Br Med J 2013 Feb 26;346:f540 [FREE Full text] [doi: 10.1136/bmj.f540] [Medline: 23444423]
33. Klein JP, Berger T, Schröder J, Späth C, Meyer B, Caspar F, et al. Effects of a psychological internet intervention in the treatment of mild to moderate depressive symptoms: results of the EVIDENT study, a randomized controlled trial. Psychother Psychosom 2016;85(4):218-228. [doi: 10.1159/000445355] [Medline: 27230863]

34. Dozois DJ, Westra HA, Collins KA, Fung TS, Garry JK. Stages of change in anxiety: psychometric properties of the University of Rhode Island change assessment (URICA) scale. Behav Res Ther 2004 Jun;42(6):711-729. [doi: 10.1016/S0005-7967(03)00193-1] [Medline: 15081886]
35. Faul F, Erdfelder E, Lang A, Buchner AG. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behav Res Methods 2007 May;39(2):175-191. [doi: 10.3758/BF03193146] [Medline: 17695343]
36. Urbaniak G, Plous S. Research Randomizer. 2018. URL: https://www.randomizer.org/ [accessed 2018-04-23]
37. Hollon SD, Thase ME, Markowitz JC. Treatment and prevention of depression. Psychol Sci Public Interest 2002 Nov;3(2):39-77. [doi: 10.1111/1529-1006.00008] [Medline: 26151569]
38. Cuijpers P, van Straten A, Warmerdam L. Behavioral activation treatments of depression: a meta-analysis. Clin Psychol Rev 2007 Apr;27(3):318-326. [doi: 10.1016/j.cpr.2006.11.001] [Medline: 17184887]
39. Hofmann SG, Sawyer AT, Witt AA, Oh D. The effect of mindfulness-based therapy on anxiety and depression: a meta-analytic review. J Consult Clin Psychol 2010 Apr;78(2):169-183 [FREE Full text] [doi: 10.1037/a0018555] [Medline: 20350028]
40. Khoury B, Lecomte T, Fortin G, Masse M, Therien P, Bouchard V, et al. Mindfulness-based therapy: a comprehensive meta-analysis. Clin Psychol Rev 2013 Aug;33(6):763-771. [doi: 10.1016/j.cpr.2013.05.005] [Medline: 23796855]
41. Teasdale JD, Segal ZV, Williams JM, Ridgeway VA, Soulsby JM, Lau MA. Prevention of relapse/recurrence in major depression by mindfulness-based cognitive therapy. J Consult Clin Psychol 2000 Aug;68(4):615-623. [doi: 10.1037/0022-006X.68.4.615] [Medline: 10965637]
42. Segrin C. Social skills deficits associated with depression. Clin Psychol Rev 2000 Apr;20(3):379-403. [doi: 10.1016/S0272-7358(98)00104-4] [Medline: 10779900]
43. Franzen PL, Buysse DJ. Sleep disturbances and depression: risk relationships for subsequent depression and therapeutic implications. Dialogues Clin Neurosci 2008;10(4):473-481 [FREE Full text] [Medline: 19170404]
44. Murphy MJ, Peterson MJ. Sleep disturbances in depression. Sleep Med Clin 2015 Mar;10(1):17-23 [FREE Full text] [doi: 10.1016/j.jsmc.2014.11.009] [Medline: 26055669]
45. Ellis A. Reason And Emotion In Psychotherapy. New York, United States: Citadel; 1994.
46. Rüdiger H, Pfingsten U. Gruppentraining Sozialer Kompetenzen GSK. Weinheim, Germany: Beltz Verlag; 2015.
47. Kaluza G. Stressbewältigung: Trainingsmanual zur Psychologischen Gesundheitsförderung. Berlin, Heidelberg: Springer-Verlag; 2015.
48. Westermann S, Moritz S, Berger T. Internet- und mobilbasierte interventionen bei schizophrenie. Verhaltenstherapie 2017 Aug 18;27(3):181-189. [doi: 10.1159/000479335]
49. Beck AT, Steer RA, Brown GK. Manual for the Beck Depression Inventory-II. San Antonio, TX: Psychological Corporation; 1996.
50. Smarr KL, Keefer AL. Measures of depression and depressive symptoms: Beck Depression Inventory-II (BDI-II), Center for Epidemiologic Studies Depression Scale (CES-D), Geriatric Depression Scale (GDS), Hospital Anxiety and Depression Scale (HADS), and Patient Health Questionnaire-9 (PHQ-9). Arthritis Care Res (Hoboken) 2011 Nov;63(Suppl 11):S454-S466 [FREE Full text] [doi: 10.1002/acr.20556] [Medline: 22588766]
51. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001 Sep;16(9):606-613 [FREE Full text] [doi: 10.1046/j.1525-1497.2001.016009606.x] [Medline: 11556941]
52. Rosenberg M. York University. 1965. Rosenberg Self-Esteem Scale (RSE) URL: https://www.yorku.ca/rokada/psyctest/rosenbrg.pdf [accessed 2019-08-16]
53. Group TW. Development of the World Health Organization WHOQOL-BREF quality of life assessment. The WHOQOL group. Psychol Med 1998 May;28(3):551-558. [doi: 10.1017/S0033291798006667] [Medline: 9626712]
54. Skevington SM, Lotfy M, O'Connell KA, WHOQOL Group. The World Health Organization's WHOQOL-BREF quality of life assessment: psychometric properties and results of the international field trial. A report from the WHOQOL group. Qual Life Res 2004 Mar;13(2):299-310. [doi: 10.1023/B:QURE.0000018486.91360.00] [Medline: 15085902]
55. Schmidt J, Lamprecht F, Wittmann WW. [Satisfaction with inpatient management. Development of a questionnaire and initial validity studies]. Psychother Psychosom Med Psychol 1989 Jul;39(7):248-255. [Medline: 2762479]
56. Ranganathan P, Pramesh C, Aggarwal R. Common pitfalls in statistical analysis: absolute risk reduction, relative risk reduction, and number needed to treat. Perspect Clin Res 2016;7(1):51-53 [FREE Full text] [doi: 10.4103/2229-3485.173773] [Medline: 26952180]
57. Vickers AJ, Altman DG. Statistics notes: analysing controlled trials with baseline and follow up measurements. Br Med J 2001 Nov 10;323(7321):1123-1124 [FREE Full text] [doi: 10.1136/bmj.323.7321.1123] [Medline: 11701584]
58. Hayes AF. Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach. New York: Guildford Press; 2013.
59. Schmidt J, Wittmann WW. Fragebogen zur Messung der Patientenzufriedenheit. In: Diagnostische Verfahren in der Psychotherapie. Göttingen: Hogrefe; 2002:392-396.
60. Burcusa SL, Iacono WG. Risk for recurrence in depression. Clin Psychol Rev 2007 Dec;27(8):959-985 [FREE Full text] [doi: 10.1016/j.cpr.2007.02.005] [Medline: 17448579]
61. Moritz S, Schilling L, Hauschildt M, Schröder J, Treszl A. A randomized controlled trial of internet-based therapy in depression. Behav Res Ther 2012 Aug;50(7-8):513-521. [doi: 10.1016/j.brat.2012.04.006] [Medline: 22677231]

62. Erbe D, Eichert HC, Riper H, Ebert DD. Blending face-to-face and internet-based interventions for the treatment of mental disorders in adults: systematic review. J Med Internet Res 2017 Sep 15;19(9):e306 [FREE Full text] [doi: 10.2196/jmir.6588] [Medline: 28916506]
63. Wentzel J, van der Vaart R, Bohlmeijer ET, van Gemert-Pijnen JE. Mixing online and face-to-face therapy: how to benefit from blended care in mental health care. JMIR Ment Health 2016 Feb 9;3(1):e9 [FREE Full text] [doi: 10.2196/mental.4534] [Medline: 26860537]
64. Berger T, Krieger T, Sude K, Meyer B, Maercker A. Evaluating an e-mental health program ("deprexis") as adjunctive treatment tool in psychotherapy for depression: Results of a pragmatic randomized controlled trial. J Affect Disord 2018 Feb;227:455-462. [doi: 10.1016/j.jad.2017.11.021] [Medline: 29154168]
65. Culpepper L, Muskin PR, Stahl SM. Major depressive disorder: understanding the significance of residual symptoms and balancing efficacy with tolerability. Am J Med 2015 Sep;128(9 Suppl):S1-15. [doi: 10.1016/j.amjmed.2015.07.001] [Medline: 26337210]
66. Israel JA. The impact of residual symptoms in major depression. Pharmaceuticals (Basel) 2010 Aug 3;3(8):2426-2440 [FREE Full text] [doi: 10.3390/ph3082426] [Medline: 27713362]
67. Constantino MJ, Arnkoff DB, Glass CR, Ametrano RM, Smith JZ. Expectations. J Clin Psychol 2011 Feb;67(2):184-192. [doi: 10.1002/jclp.20754] [Medline: 21128304]
68. Greenberg RP, Constantino MJ, Bruce N. Are patient expectations still relevant for psychotherapy process and outcome? Clin Psychol Rev 2006 Oct;26(6):657-678. [doi: 10.1016/j.cpr.2005.03.002] [Medline: 15908088]
69. Scogin FR, Hanson A, Welsh D. Self-administered treatment in stepped-care models of depression treatment. J Clin Psychol 2003 Mar;59(3):341-349. [doi: 10.1002/jclp.10133] [Medline: 12579549]
70. van Straten A, Seekles W, van't Veer-Tazelaar NJ, Beekman AT, Cuijpers P. Stepped care for depression in primary care: what should be offered and how? Med J Aust 2010 Jun 7;192(11 Suppl):S36-S39. [Medline: 20528706]
71. Lambert M, Karow A, Deister A, Ohm G, Bargel S, Peter H, et al. RECOVER: modell der sektorenübergreifend-koordinierten, schweregrad-gestuften, evidenzbasierten versorgung psychischer erkrankungen. In: Amelung VE, Eble S, Hildebrandt H, Knieps F, Lägel R, Ozegowski S, et al, editors. Innovationsfonds: Impulse Für Das Deutsche Gesundheitssystem. Hannover: Medizinisch Wissenschaftliche Verlagsgesellschaft; 2017:252-265.
72. Salomonsson S, Santoft F, Lindsäter E, Ejeby K, Ljótsson B, Öst LG, et al. Stepped care in primary care - guided self-help and face-to-face cognitive behavioural therapy for common mental disorders: a randomized controlled trial. Psychol Med 2018 Jul;48(10):1644-1654. [doi: 10.1017/S0033291717003129] [Medline: 29095133]
73. Karyotaki E, Kemmeren L, Riper H, Twisk J, Hoogendoorn A, Kleiboer A, et al. Is self-guided internet-based cognitive behavioural therapy (iCBT) harmful? An individual participant data meta-analysis. Psychol Med 2018 Nov;48(15):2456-2466 [FREE Full text] [doi: 10.1017/S0033291718000648] [Medline: 29540243]
74. Donkin L, Christensen H, Naismith SL, Neal B, Hickie IB, Glozier N. A systematic review of the impact of adherence on the effectiveness of e-therapies. J Med Internet Res 2011 Aug 5;13(3):e52 [FREE Full text] [doi: 10.2196/jmir.1772] [Medline: 21821503]
75. Jelinek L, Hauschildt M, Wittekind CE, Schneider BC, Kriston L, Moritz S. Efficacy of metacognitive training for depression: a randomized controlled trial. Psychother Psychosom 2016;85(4):231-234. [doi: 10.1159/000443699] [Medline: 27230865]
76. Teasdale J, Scott J, Moore R, Hayhurst H, Pope M, Paykel E. How does cognitive therapy prevent relapse in residual depression? Evidence from a controlled trial. J Consult Clin Psychol 2001 Jun;69(3):347-357. [doi: 10.1037//0022-006X.69.3.347] [Medline: 11495165]
77. Titov N, Dear BF, Johnston L, Lorian C, Zou J, Wootton B, et al. Improving adherence and clinical outcomes in self-guided internet treatment for anxiety and depression: randomised controlled trial. PLoS One 2013;8(7):e62873 [FREE Full text] [doi: 10.1371/journal.pone.0062873] [Medline: 23843932]

Abbreviations

ABC (-protocol): activating event, beliefs, consequence
ANCOVA: analysis of covariances
BDI-II: Beck Depression Inventory-II
CAU: care-as-usual
CBT: cognitive behavioral therapy
ITT: intention-to-treat
MD: major depression
PHQ-9: Patient Health Questionnaire–9
PP: per-protocol
RCT: randomized controlled trial
RSE: Rosenberg Self-Esteem
URICA: University of Rhode Island Change Assessment
WHO: World Health Organization

WHOQOL-BREF: World Health Organization Quality of Life–abbreviated version
ZUF-8: Fragebogen zur Patientenzufriedenheit (patient satisfaction questionnaire)

Edited by G Eysenbach; submitted 02.04.19; peer-reviewed by A González-Robles, A D?az-Garc, A Baglione, F Lanfranchi; comments to author 22.05.19; revised version received 19.06.19; accepted 19.07.19; published 02.10.19.

Please cite as:
Bücker L, Schnakenberg P, Karyotaki E, Moritz S, Westermann S
Diminishing Effects After Recurrent Use of Self-Guided Internet-Based Interventions in Depression: Randomized Controlled Trial
J Med Internet Res 2019;21(10):e14240
URL: https://www.jmir.org/2019/10/e14240
doi:10.2196/14240
PMID:31579014

Original Paper

Promoting Social Connection and Deepening Relations Among Older Adults: Design and Qualitative Evaluation of Media Parcels

Isabela Zaine[1*], BSc, MSc, PhD; David Mark Frohlich[2*], BA, PhD; Kamila Rios Da Hora Rodrigues[3*], BSc, MSc, PhD; Bruna Carolina Rodrigues Cunha[1,4*], BSc, MSc, PhD; Alex Fernando Orlando[1*], BSc, MSc; Leonardo Fernandes Scalco[1*], BSc; Maria Da Graça Campos Pimentel[1*], BSc, MSc, PhD

[1]University of São Paulo, Institute of Mathematics and Computer Sciences, São Carlos - SP, Brazil

[2]University of Surrey, Department of Music and Media, Guildford, United Kingdom

[3]Federal University of São Carlos, Department of Computer Sciences, São Carlos, Brazil

[4]Federal Institute of Education, Science and Technology of São Paulo, Capivari, Brazil

[*]all authors contributed equally

Corresponding Author:
Isabela Zaine, BSc, MSc, PhD
University of São Paulo
Institute of Mathematics and Computer Sciences
Avenida Trabalhador São Carlense, 400 - Parque Arnold Schimidt
São Carlos - SP, 13566-590
Brazil
Phone: 55 163373 9700
Email: isabela.zaine@gmail.com

Abstract

Background: Being socially connected is related to well-being, and one way of avoiding social isolation is to deepen existing relationships. Even though existing relationships can be reinforced by regular and meaningful communication, state-of-the-art communication technologies alone do not increase the quality of social connections. Thus, there is a need for the involvement of a trained human facilitator in a network of older adults, preferably for a short period, to promote the deepening of their relationships.

Objective: This study aimed to evaluate the hypothesis that a human-facilitated, media-sharing social networking system can improve social connection in a small group of older people, who are more vulnerable to social isolation than most, and deepen their relationships over a period of a few weeks.

Methods: We conducted the design and evaluation of *Media Parcels*, a novel human-facilitated social networking system. *Media Parcels* is based on the metaphor of a facilitator collecting and delivering parcels in the physical mail. Extending the metaphor, the system supports a facilitator in designing time-based dialogue requesting parcels from participants that bring out their memories and feelings, in collecting the parcels, wrapping them in annotations that communicate the corresponding requests, and delivering the wrapped parcel to a target person. Qualitative evaluation was carried out in two trials with a group of three people each, one with family members (children and father; aged 55, 56, and 82 years old) and the other with a group of friends (aged 72, 72, and 74 years old), over two weeks. In each trial, data were collected in three interviews (pre-, mid-, and posttrial) and via system logging.

Results: Collected data indicate positive social effects for deepening and developing relationships. The parcel metaphor was easily understood and the computational system was readily adopted. Preferences with regard to media production or consumption varied among participants. In the family group, children preferred receiving media parcels (because of their sentimental value) to producing them, whereas the father enjoyed both. In the friendship group, preferences varied: one friend enjoyed both producing and receiving, while the other two preferred one over the other. In general, participants reported a preference for the production of items of a certain type depending on the associated content. Apart from having a strong engagement with the system, participants reported feeling closer to each other than usual.

Conclusions: For both groups, *Media Parcels* was effective in promoting media sharing and social connections, resulting in the deepening of existing relationships. Its design informs researchers who are attempting to promote social connection in older adults.

(J Med Internet Res 2019;21(10):e14112) doi:10.2196/14112

KEYWORDS

social interaction; interpersonal relations; communication; intervention; experience sampling; mobile apps; photography; video-audio media; elderly

Introduction

Background

The concept of connectedness is multidimensional and there are multiple definitions and types of connectedness, including social connectedness, that focus on interpersonal relationships. Social connectedness, as defined by Lee and Robbins [1,2], includes feelings of being in a close relationship with the social world, which is critical to one's sense of belonging and is based on proximal and distal relationships. As social interactions may be transient, feelings of being connected to others may change. Thus, social connectedness may also correspond to short-term experiences of belonging [3]. Having regular and meaningful contact with others usually increases social connectedness, but this type of contact may be hindering for older people because of mobility and geographic constraints.

The global population aged 60 years old or above is growing at a rate of about 3% per year, which is faster than all younger age groups. At present, in Europe 25% of the population is already aged above 60 years old, and in Latin America they account for 12% of the population [4]. This shift in the world's demographics has led to a need for action across multiple sectors to enable older people to age well and remain a resource to their families, communities, and economies. Thus, successful ageing and healthy ageing have been topics of interest in some of the most recent studies conducted among the older population. The World Health Organization [5] defines successful ageing as the process of developing and maintaining the functional ability that enables well-being in older age. By including well-being in this definition, it goes beyond physical health as it includes domains such as happiness, satisfaction, fulfillment, and feelings of belongingness.

Older people often emphasize the role of social integration and well-being associated with successful ageing [6]. As they age, they wish to maintain their identities and social roles, have relationships, remain autonomous, feel safe, feel like they still have the potential for personal growth, and be able to enjoy life [7-9].

Older adults are more vulnerable to social isolation than the rest of the population. Typically, their social networks shrink with age because of bereavement, relocation, and reduced mobility [10]. Moreover, older adults who are socially isolated have been shown to be more depressed and disabled, be in poorer health, and report lower levels of well-being than those who are socially connected [11]. Thus, it is important that they are able to maintain and deepen social connections with their remaining family and friends and to establish new ones.

Maintaining regular social connections can be challenging for older people living alone, who may struggle to travel to meet distant contacts in person or attend in-person social events. Web-based contact seems to be an obvious solution, particularly given the rise in internet access via mobile devices for adults aged 65 years or older in the United Kingdom [12], Brazil [13], and the United States [14], among others. Such technology can support older adults [15] who prefer traditional one-to-one channels of communication [16,17].

Social connectedness can be stimulated by experiences that remind people of their social relationships. For example, looking at a photograph or listening to a song may remind them of a loved one and their relationship with them. Thus, feelings of social connectedness do not necessarily require physical proximity. Based on these insights, our hypothesis is that a human-facilitated, media-sharing social networking system can improve social connection in older people and deepen their relationships upon deployment over a few weeks.

To evaluate our hypothesis, we designed a new type of social networking system called *Media Parcels* to address social connection in older people and deepen their relationships over a few weeks with the mediation of a human facilitator. Media collection and sharing, orchestrated by the facilitator, lies at the heart of the system, as this activity is meant to stimulate user reflection about current relationships, both strong and weak. In the *Media Parcels* system, parcels of media are solicited by the facilitator for later consumption by specific recipients, with the types of media requested designed to encourage reciprocal intimacy and self-disclosure between an older person and selected social contacts.

Both the intervention duration and the group size are important in the proposed design. One of the main reasons is that a large group size would imply developing a large number of relationships, which is difficult to achieve in a short period and challenging for a facilitator to coordinate. Second, a duration of a few weeks allows the involvement of the facilitator to be short term, and finally, this would allow the study to take advantage of the temporary novelty effect observed with new technology adoption.

In the current design of the system, media collection and distribution are done by a human facilitator among 3 users. In future designs, the facilitator might be supported or eventually replaced by an algorithm designed to allow larger networks.

Social Networking and Older People

Much research exists on the value of sharing digital media and other physical memorabilia for maintaining relationships but not the honesty and intent [18-21]. Here, we will concentrate on more recent broad social media services and studies of media sharing through social networking systems, especially those targeted at the older population.

Facebook use is now part of daily life for many people around the globe, and a huge number of posts made by friends are delivered continuously to users. An early study on self-disclosure on Facebook by Park et al [22] with 317 student

participants showed that the number of and positivity of posts is associated with the intimacy of the relationship, but not the honesty and intent. In a later study with 243 participants, Orben and Dunbar [23] investigated how passive consumption of personal disclosures affects the development of relationships. Passive consumption occurs when a person reads the posts of others without directly interacting with them. Their results suggest that reading high intimacy, self-disclosures increases relationship closeness, triggering similar relationship formation in real-life interactions.

When focused especially on the elderly population, the use of social media is slightly different to other groups. Rebelo [24] conducted an exploratory study in Portugal with 4 older adults to understand the motivations and interests of the elderly population when using social networks (in particular Facebook). The researcher found that, for this particular group, motivation was mainly related to solitude, belonging, reunions, and willingness to improve intergenerational relationships. Regarding shared content, the elderly stated that they like discussing memories but were concerned about privacy, and they thought that most of the published content on the social network site was inadequate or uninteresting. Chakraborty et al [25] studied the privacy-preserving behaviors of older Facebook users in direct contrast to their self-disclosure. They analyzed the profile information and privacy settings of 134 Facebook users aged above 55 years, together with 50 of their friends, and they observed that older adults hide or share information on their profiles depending on what information their contacts share. They tend to be more conservative about information sharing, in line with the findings of Lindley et al [16] and Pedell et al [17].

Sinclair and Grieve [26], in turn, investigated whether older adults could derive social connectedness from Facebook and whether the levels of social connectedness were similar to those seen in younger samples. The analysis revealed that Facebook social connectedness emerged as a separate factor to offline social connectedness, with correlations between the factors indicating that they were distinct constructs. The participants reported levels of Facebook-derived social connectedness similar to those seen in younger samples. About the social effects and benefits caused by the networks, Quinn [27] examined the engagement among novice social media users, aged 65 years and older, in 4 cognitive domains: attention, processing speed, working memory, and inhibitory control. Results reported include the improvement of intervention participants in inhibitory control. Quinn argued that the findings demonstrated that the benefits of social media use at older ages extended beyond mere social engagement and into other domains of everyday well-being.

Chen and Schlz [28] conducted a systematic review to verify the effect of information communication technology on elderly social isolation. Their findings suggested that although the technology lessened social isolation, the technology alone does not guarantee quality of communication among older adults. Furthermore, when communication is not reciprocal, technology use may increase social isolation [29].

Several novel media sharing systems have been designed for the older population, with differing levels of effects on relationships. A number of these systems use situated displays in the home to share materials at a distance, such as photographs, text messages and broadcast media [30-33]. For instance, Garattini et al [34] designed a system to promote opportunistic social interaction among elderly people. Using a tablet, the system broadcast news of different topics and presents functionalities to enable group conversations through voice and text. The researchers conducted a 10-week study with 19 elders and some of their friends and family members. The results suggested that although the broadcasts encouraged social interactions, the quality of the engagement was limited by the absence of an approach to share personal information to help users become familiar with one another.

Waycott et al [31] developed a tool to facilitate message and media sharing and conducted a study in which caregivers used the tool to communicate with their elderly clients. Their results showed that photographs with captions were able to increase and enhance communication, and were well-suited to promote psychosocial care. Similarly, Abrahão et al [35] conducted a mobile digital storytelling study in a care home for older people. The creation of the stories was facilitated by either a formal or informal carer and focused mainly on the resident, capturing aspects of everyday life such as visits, social events, therapy sessions, and health reports. The results showed that the technology stimulated expressivity and creativity in the resident, as well as richer conversations between the resident and other people.

Cornejo et al [32] developed a system geared toward enhancing older people's offline interactions with their family. The results emphasized how information shared on social media could provide conversational context for the elderly, prevent isolation, and increase offline conversations.

Most studies report positive benefits for maintaining or making new relationships through lightweight multimedia messages. Given the established relationship between the level of self-disclosure and relationship closeness both online and offline, there is an opportunity to explicitly encourage media exchange related to the existing relationship. By encouraging self-disclosure, we mean to encourage participants to talk about feelings and emotions, to talk about topics they do not regularly talk about (eg, intimate things), and to talk about themselves and their relationship toward each other so that using an app might be easier than saying it face-to-face.

Methods

Media Parcels Design

Media Parcels is a novel social networking system designed to promote facilitated media exchange between users. The interaction underlying *Media Parcels* is based on the metaphor of parcel delivery in the physical mail. First, a facilitator, upon specific requests to participants, collects media and wraps them in text commentary, bringing out their memories and meaning. Next, the facilitator passes the wrapped media parcel to a target person, who in turn unwraps them.

In the general case, Media Parcels supports facilitators in orchestrating interactions among any number of users by intervening in 2 steps: in phase 1 by requesting media parcels from a user and, in phase 2 by sending the parcels to any number of (other) users.

The particular case of deploying Media Parcels for the balanced interaction among 3 users (P1, P2, and P3) orchestrated by a facilitator is illustrated in Figure 1. For the same scenario, Table 1 details the request of parcels of media (from-to).

As currently designed, *Media Parcels* relies on a facilitator in charge of managing the interaction for the duration of the intervention. In a nonfacilitated approach, *Media Parcels* could, in principle, pass parcels of media between multiple providers and recipients of media, at various times, *ad infinitum*.

In contrast to most current social media systems that rely on the spontaneous posting of media for feedback, the *Media Parcels* approach is based on directed requests for media items which are then shared within a group.

For the deployment of *Media Parcels*, we used the Experience Sampling and Programmed Intervention Method (ESPIM) and an associated platform that support specialists (eg, health professionals) in communicating with their users at particular times of the day via a mobile app [36,37]. ESPIM is inspired by the experience sampling method for collecting self-reports in psychology proposed by Csikszentmihalyi et al [38], combined with the concept of programmed instruction [39]. The ESPIM software platform was designed to support health care and learning interventions in their natural environments throughout the day [36]. Such interventions can comprise open-ended and multiple choice questions, media requests, and deliveries. While creating the interventions, the professional defines the time-based moments in which they should be prompted on the user's device [40].

Figure 1. Media Parcels exchange among three persons (P1-P3) in a two-phase interaction orchestrated by a Facilitator, who "wraps" the parcel by adding commentaries.

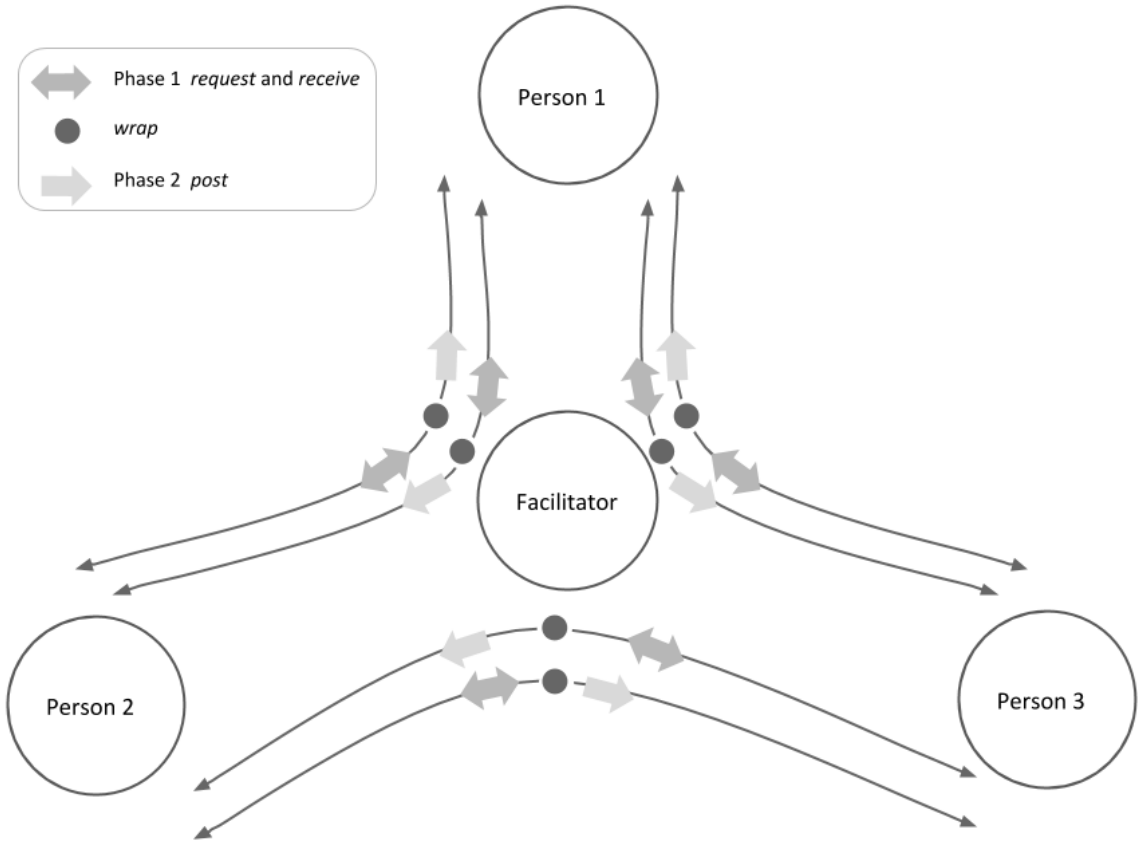


Table 1. Operations for *Media Parcels* exchange among persons P1, P2, and P3 in request-wrap-post interaction orchestrated by a facilitator.

Phase and facilitator (role)	P1	P2	P3
Phase 1			
Media parcel (Request)	from (P1) to (P2)	from (P2) to (P1)	from (P3) to (P1)
Media parcel (Receive)	from (P1) to (P3)	from (P2) to (P3)	from (P3) to (P2)
Wrapping			
Media parcel (Wrap)	—[a]	—	—
Phase 2			
Media parcel (Post)	from (P2) to (P1)	from (P1) to (P2)	from (P1) to (P3)
Media parcel (Post)	from (P3) to (P1)	from (P3) to (P2)	from (P2) to (P3)

[a]Not applicable.

Approach Toward Social Connection and Deepening Relations

Aiming at supporting a facilitator who promotes social connection and deepening of relationships in a small network, comprising older persons and their social contacts, content exchanged via Media Parcels should encourage self-disclosure in the scope of the social relations involved while ensuring both intimacy and privacy.

Facilitator

The role of the facilitator includes: (1) designing media requests associated with feelings that encourage self-disclosure and reflection over a specific social contact; (2) wrapping together the media elements received and the associated feelings, without editing them; and (3) mediating the interaction between a pair of users by sharing the content produced between both ends of the targeted relationship.

For the purposes of this study, such a role demands professional skills, so in the case of the two trials described here the facilitator had a background in clinical psychology.

Media Requests

Requests for media were deliberately designed to be thought provoking and vehicles for self-reflection and disclosure. For example, the following requests and questions were used, among others, to elicit responses across a variety of media forms, personalized by name:

- Take a picture of an object that is special to you and explain why (image and text).
- Record a short audio clip of yourself singing a snippet of a song you like and explain why you like it (audio and text).
- Record a short audio to X, saying what he or she means to you (audio).
- What do you like to do with X? Why? (text).
- Send me a picture of a present X has sent you in the past. What was the occasion? How did you feel about it? (image and text).
- Record a short video of you saying something X often says (video).

Media requests were inspired by methods used in design and behavioral psychotherapy. We used two strategies for the design of media requests, one concerning the type of media, and the other, the media content. Different types of media may produce different emotional responses [41], so we designed media requests to be balanced across image, audio, text, and video to provide participants with all types of media once, and regarding media content, the media requests were designed to probe participants' relationships and themselves.

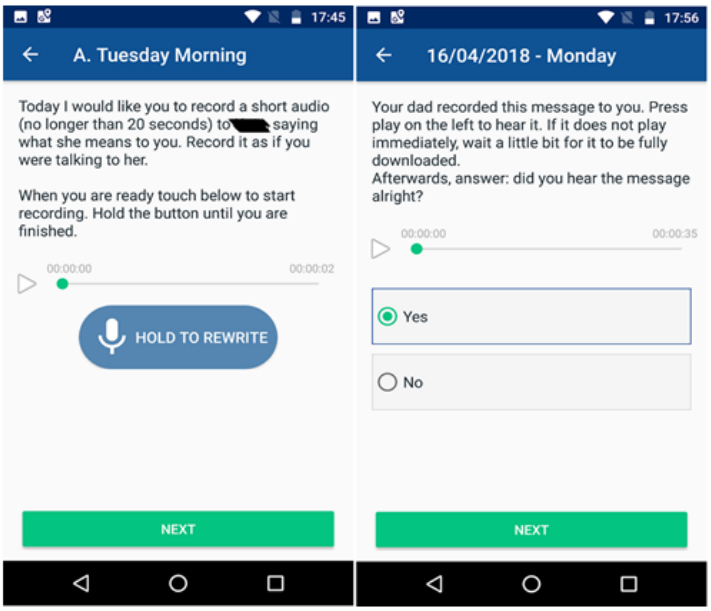
Behavioral psychological interventions make use of asking questions, making requests, and giving instructions or suggestions for clients to carry out specific or generic actions outside the therapeutic setting [42]. Asking questions about happenings or feelings serves to gather information and also promotes self-observation, self-reflection, and self-knowledge. In our study, the media requests and questions about participants' feelings were designed to encourage all participants to reflect on their relationships and express their feelings toward one another by producing media with emotional content. Furthermore, specifically for the older participants, the questions and requests also encouraged reminiscence of precious moments and self-disclosure.

The questions and requests might also be referred to as relationship probes, as they have the character of cultural probe questions for self-report except they are focused on personal traits and relationships rather than culture [43]. Cultural probes is a technique used to inspire ideas in a design process. Typically, the probes are small packages that can include any sort of artifact along with evocative tasks that allow participants to keep record of specific events, feelings, or interactions. It serves as a means of gathering inspirational data about people's lives, values, and thoughts.

Media Wrapping and Sharing

For each media item collected, before sending the item to the target person the facilitator included a text comment to expose the request that created it. An example is given in Figure 2, with the facilitator requesting for one person to record a piece of audio for another person saying what she means to you. When sharing that media, the facilitator includes "Your dad recorded this message to you". As indicated by this example, no editing relative to the media collected took place.

Figure 2. Typical media request and delivery screens. Left: media request in phase 1 to the older adult. Right: media delivered to another participant in phase 2.



Older Person's Social Contacts

Social contacts of an older person may be inter- or intragenerational, such as younger family members (their children or grandchildren) for intergenerational contacts, and the same generation friends or family members for intragenerational contacts. According to Lindley et al [16], the dynamics of family and friend relationships are very different, especially in terms of expected social support. Considering the two different types of social contacts, we designed two trials devoted to each type of relationship: one older person with two younger family members, and three same-generation older individuals that were friends.

For practical reasons related to our international collaboration, we conducted the first trial with family participants in the United Kingdom and the second trial with friendship participants in Brazil. The system and methods for both trials were largely the same and are reported in study 1 for the UK context, with variations outlined in study 2 for the Brazilian context.

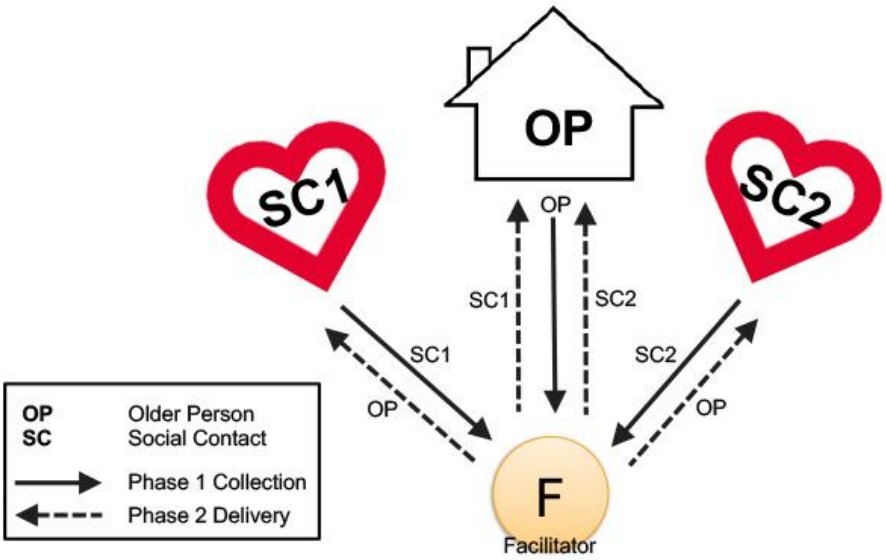
The studies reported in this paper were designed so that phases 1 and 2 lasted around one week each. This amounted to a fixed period intervention aimed at increasing social connection between these partners. We designed a media collection phase (phase 1) and a media sharing phase (phase 2), so that participants could clearly discriminate between producing and receiving personal media content. In both studies, participants were aware of the two distinct phases and that the produced media in phase 1 would not be immediately shared with the other participants. They were also informed that the human facilitator, who was a clinical psychologist, would be in charge of sharing the produced media and its associated feelings from phase 1 with the other participants on phase 2, without editing the content. Participants were also aware that, once the facilitator mediated the media sharing, this person would have access to the content of the media and its associated feelings. In addition, the participants were informed that, after phase 2, the facilitator would not mediate their interaction anymore; thus, if they wished to talk about the media content received, they should use their conventional communication channels.

Study 1: Family Trial

This study focused on supporting an older person through media exchanges with two younger family members. The exchange of media parcels was asymmetric because they were not collected and delivered between the younger participants.

Figure 3 shows the network configuration for study 1. An older adult who lives alone exchanged media (text, audio, image, and video) with 2 other social contacts. The facilitator (the first author) orchestrated the exchange in 2 phases. In phase 1, lasting a week, media relating to each pairwise relationship with the older person was collected by the facilitator at regular intervals. In phase 2, also lasting a week, that media was distributed to the reciprocal partner in each pair at regular intervals. As the older person is linked with 2 people, while those people are only linked to one and not each other, we collected and distributed twice as much content from the older person as from the reciprocal partners, as shown by the arrows in Figure 3.

Figure 3. Network configuration and media flow in the Media Parcels system for study 1.



Media

A typical media collection and sharing screen is shown in Figure 2, as previously presented. Requests for media were deliberately designed to be thought provoking and vehicles for self-reflection and disclosure. For example, the following requests and questions were used, among others, to elicit responses across a variety of media forms, personalized by name:

- Take a picture of an object that is special to you and explain why (image and text).
- Record a short audio clip of yourself singing a snippet of a song you like and explain why you like it (audio and text).
- Record a short audio to X, saying what he or she means to you (audio).
- What do you like to do with X? Why? (text).
- Send me a picture of a present X has sent you in the past. What was the occasion? How did you feel about it? (image and text).
- Record a short video of you saying something X often says (video).

Participants

Recruitment requirements for the older person were as follows: (1) aged 60 years or older; (2) lives alone; (3) is based in the Guildford area, United Kingdom; and (4) is able to nominate at least two other social contacts who are able to participate. All names were changed to preserve participants' identities. We formally broadcasted the search for participants for 6 weeks through the University of Surrey academic email to the existing network of research groups and personal contacts, on the research group website, and by partnering with Age UK Surrey and the University of the Third Age, Guildford. The participants recruited were resultant of personal contact.

Participants were 1 older adult, Paul (aged 82 years), and 2 of his children: Karen (aged 55 years) and Charles (aged 56 years). Paul chose them because he considered them relevant people for social support in his life. Paul is an electrical and electronic engineer and still works part-time, mentoring engineers at a company. He did not have any cognitive impairment according to the Mini-Mental State Examination to measure cognitive health [44]. He was also living alone for the past 6 months and feeling more isolated since his wife passed away. Charles, his son, lives close to him. His daughter, Karen, lives in Kenya, Africa, but was visiting her father for the first week of the trial. The field data collection was done from late April 2018 to late May 2018.

Pretrial Interview

A pretrial semistructured interview was conducted face-to-face with each participant. With the older person, we enquired about his relationship with his two children. To access feelings of a specific social connection at the individual level between Paul and his children, we designed a Relationship Semantic Differential Scale (RSDS, Multimedia Appendix 1) with 16 pairs of contrasting attributes related to social relationships. Participants were instructed to express their agreement with the attributes on a 7-point scale. As typical of semantic differential scales, the closer the participant responds to one of the attributes the more they agree with it [45]. Paul was asked to rate his relationship with each child separately. Finally, the *Media Parcels* mobile app was downloaded and installed on a dedicated tablet. Paul was given this tablet and trained in how to use it. This took a total of about 1 hour and 30 min.

Each of Paul's children also answered a similar semistructured interview enquiring about his or her relationship with their father and responded to the RSDS. The *Media Parcels* mobile app was downloaded to their existing smartphones and they were trained in how to use it. This took about 30 min per person.

Trial Period

Following the pretrial interview, the system was activated on all 3 devices and proceeded to collect media parcels from participants over 7 consecutive days in the first week of the trial. These were analyzed and selected by the facilitator for

redistribution over the following 8 consecutive days of the second week for the social contacts and 9 consecutive days for the older person. The amount of media parcels received in the second week was based on the amount of relevant media parcels produced by participants in the first week; therefore, it varied among participants. The older person, Paul, received 3 collection messages or 1 delivery message from the system each day, whereas his children Karen and Charles received only 1 collection or 1 delivery message a day. Notifications of each message sounded and appeared on the receiving devices, but users were able to defer answering or reading them until later. Users could also choose to decline answering any request. Once a media delivery had been seen, users could not directly respond with feedback to the person who sent the media, as is usual on most social media platforms. It was a design decision not to support interactional closure to encourage participants to contact each other offline or using their existing preferred communication channels. This decision was based on the fact that new technology solutions on communication mostly force users to abandon their usual channels to adopt new ones. This can be aversive to those people who take more time learning to use a novel technology, which is the case for most of the older population [46].

Midtrial Interview

At the end of the first week (phase 1) of the trial, we conducted a midtrial interview. Participants once again each responded to the RSDS and were interviewed about the feelings generated by producing materials about their relationship. Participants were also briefed about the delivery of phase 2 of the study.

Posttrial Interview

After phase 2, we conducted final semistructured interviews with all participants, separately, about their relationships, connectedness toward one another, and experience using the system. They then responded to the RSDS. The meetings lasted an average of 30 min for each social contact and 1 hour for the older person. In this last interview, they also evaluated their experience with the ESPIM mobile app by answering a User Experience Questionnaire (UEQ) [47] and a System Usability Scale (SUS) [48]. All interviews were conducted in the participants' native language, English.

Study 2: Friendship Trial

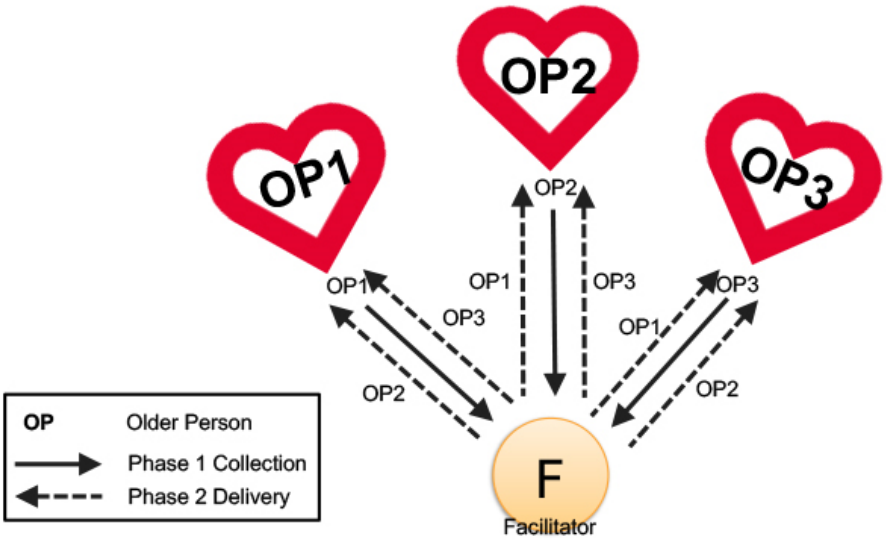
To test out the value of the system for deepening relationships with same generation friends, we recruited 3 older people with different levels of acquaintance, as described below. We decided to treat all these participants equally in the trial by using symmetrical media sharing between all three, where P1, P2, and P3 are three elderly ladies. This means the scripted dialogue between participants and the facilitator implies that media were collected in phase 1 from each participant about their other 2 partners and distributed to those partners in phase 2.

Participants

Recruitment requirements were as follows: (1) aged 60 years or older; (2) lives alone; (3) is based in São Carlos, Brazil; and (4) is able to nominate at least two other same-generation social contacts to also participate. The research was broadcast in a digital literacy course for older adults. In total, three elderly ladies volunteered. They described themselves as friends or acquaintances. Ronda is a 72-year-old retired teacher who was divorced and had a son. Linda is also aged 72 years and is a retired administrative assistant) who was never married and has had a boyfriend for the past 8 years. Finally, Irene is 74 and a retired laboratory assistant. All participants were Brazilian. They had some experience dealing with smartphones and tablets. None of them had any cognitive impairments.

The network configuration and media flow in study 2 is presented in Figure 4. Both Ronda and Linda were more connected to Irene, who acted as a friendship mediator between the two others. Ronda and Linda were friends for about 10 years and they both described being close to each other. Linda and Irene have been friends for over 30 years. They worked together, and before their retirement they were very close. Irene said their relationship started to go cold after a while, especially after Linda started dating. Ronda and Linda consider themselves to be friends, they have known each other for 10 years, though they do not call or arrange to meet in person as this is always mediated by Irene. In addition, most of the time Irene is the one who initiates social contact with both friends. None of them described themselves as socially isolated, but Ronda expressed interest in deepening her relationship with both her friends, Linda and Irene would like to revive their friendship that had become distant over the years, and Linda would also like to become closer to Ronda.

Figure 4. Network configuration and media flow in the Media Parcels system for study 2.



Procedure

The procedure was identical to study 1: (1) pretrial interview; (2) media collection; (3) midtrial interview; (4) media sharing; and (5) posttrial interview. In this case, the interviews were conducted in Portuguese. The total length of the data collection, from pre- to posttrial, was 20 days. Media were shared symmetrically as described above. The media requests were specially tailored to suit the particularities of the three participants and to deepen their relations. As before, requests were to share information about themselves (personal requests) and about their relationships with each other (relationship requests).

In phase 1, a total of 14 requests were designed for each participant: 7 personal and 7 about the relationships. The requests involved the same media types as in study 1 (text, audio, video, and image) as well as text commenting about their feelings toward what they had produced. Among the images requested, there were photos of people (friends participating in the study), objects (presents they received from each other, special objects, and objects that reminded them of each other), and places (favorite spot in the house). The audio requests were to share a memory from childhood, talk about a miracle in their lives, and record a song snippet. The video requests were to record an ongoing activity, talk about a good day they spent with the friends, and record a video message to the other friends. The text requests were to write about what they would like to do more with the friends and to write about their favorite singers, television shows, and activities. A total of two requests were sent out each day, one in the morning and another in the afternoon for 7 consecutive days. As in study 1, participants could respond to the requests immediately after they received the sound reminder or any other time during the day. In phase 2, we shared the most compelling media via the system, which, in the case of this study, were the media and associated feelings with more emotional and personal content. All participants received 1 media parcel per day for 9 consecutive days.

Results

Study 1: Family Trial

User Experience, Engagement, and Patterns of Interaction

The results of the user experience and usability rating scales were very positive, with the SUS results indicating an average score of 77.5 points. This is about 10 points higher than average, and a B+ on a scale ranging from A to F. The UEQ scores were at excellent or above average across all dimensions. Overall, participants were able to use the app without any issue other than difficulty in identifying the status of video recording when the Android video app is used, an issue not under the control of the designers. Such positive results, which relate to the attention given to guidelines during the design of the mobile app [40], contrast with the poor accessibility of many apps when used by older people [49,50]. Looking ahead, the experience and usability scores were very similar for study 2 (average SUS=79.17 and slightly higher UEQ scores). Therefore, the rest of the paper concentrates on the process of using the system and its value for relationships.

In phase 1, the media collection phase, all participants responded to more than half the media requests. Paul responded to 76% (16/21) of media requests on a daily basis. He interacted with the app immediately after the sound reminder in about 38% (8/21) of cases, and preferred to use the app when he had time to engage in the requests:

> *I heard the alarm, but if I was busy at that moment, I just ignored it (...). I knew I had to do the tasks, so I tried to come back to it when I had time.* [Paul]

On average, he spent 5 min per interaction, although he spent more time on the requests that required him to retrieve materials, such as pictures, written material, and objects.

Karen responded to 100% (7/7) of media requests. She interacted with the app on two separate days and preferred to respond to most of the requests in a row. It took her an average of 3 min per interaction. She never responded to the requests immediately after the sound reminder:

> *I preferred to do it all at once, in a day I had time.* [Karen]

Charles presented the lowest rate of engagement, 55% (5/9). In the intermediate interview, he said he did not interact with the app as much because he had a busy week. He spent 10 min on average in each request and also preferred to interact with the system when he had time, despite the sound reminders.

In phase 2, the media sharing phase, the engagement with the app was higher. Paul and Karen interacted with all of the media shared and Charles did not view only one of the parcels of media, interacting with 87% of them (7/8). The pattern of interaction with the app also changed in phase 2. This time, Karen and Charles interacted with the app almost daily and usually before the sound alert was triggered as a reminder. This suggests an expectation to receive the media produced by the father, which was confirmed in the posttrial interviews in which they reported being curious about what they would receive that day.

Type of Media and Media Content

The requests were effective in leading to self-disclosures revealing user's personalities, memories, and relationships. For example, a response to a request for a picture and explanation of a special object by the older person is shown in Figure 5. The story of this object is touching and something Paul may not have posted spontaneously on a social media site.

On the basis of their experiences producing and receiving media, we asked participants about their favorite types of media and media content in each phase of the study. For Paul, the most important feature was the content of the media and not the type itself. Thus, when producing media, his least favorite was typing (text) "It was boring, I am slow (...) but not impossible." Regarding the type of media received, he said that one of his favorites was a song snippet from his daughter singing ("You are my sunshine"). He said that this media inspired him to get back to learning the Ukulele, so that he could play that song for her in the future. Charles enjoyed more producing videos because he thought it was an opportunity for him to show something he was proud of. He also reported that he would have enjoyed recording longer videos. He enjoyed receiving the photos and videos best because they were visual. As for Karen, she enjoyed producing all types of media and was pleased that she had the opportunity to do a little bit of everything. She thought that the different types of media made the tasks more enjoyable and less predictable. When she received the media, she enjoyed text, audio, and video more than photos:

> *The photos were not surprising, they were not new to me (...) the others were an expression of dad in that moment.* [Karen]

Figure 5. Examples of media collected in phase 1 (image at the left and text at the right).

Feelings (text): "The cross belonged to my wife and before that to her mother. They wore it and treasured it. I feel that I have a responsibility to keep it close and safe."

Producing Media Versus Receiving Media

Both Karen and Charles preferred receiving media rather than producing them. They felt good being told good things about themselves by their father and because it was a surprise to see what their father had to say to them. Also, they considered this phase more effortless. Paul also reported that the second week

was easier, but he enjoyed both producing and receiving the media:

> *I enjoyed them (...) it was a good surprise (...) and I was proud of them [Karen and Charles] for participating (...). It was like aconversation.* [Paul]

Feelings of Specific Social Connectedness

All participants reported feeling closer to each other and contacting each other more than usual during the *Media Parcels* trial. This is reflected in the relationship (RSDS) scores taken at the beginning, middle, and end of the trial shown in Multimedia Appendix 2. Scores generally rose in the trial for Paul and Karen. Interestingly, their rise in scores was greatest in the middle of the trial, after media collection but before delivery. This suggests an effect of feeling closer to someone when encouraged to think about them through a media collection task. Charles' scores remained about the same, perhaps because he found the exercise harder than the other two in terms of expressing his feelings (see Multimedia Appendix 2).

Paul reported that he felt that his contact and relationship with his kids changed during the study:

> *We ended up being more thoughtful, because of the questions asked. We have become closer and honest to each other.* [Paul]

He reported that producing and receiving the media and talking about the feelings associated with them made him think about being more effective and also about how to get things done in general:

> *I realized I was letting things slide. You stimulated my attitude (...) because I had to do the activities it also encouraged me to do other things I needed to do.* [Paul]

He also thought that doing the activities made him think about his kids much more than usual, and he contacted them more than usual:

> *It made me think more about emotions and feelings (...) we usually let it go and don't make time for ourselves. I was happy to do that (...) I was relieved to have a reason to do that.* [Paul]

Charles reported that, for him, talking about feelings was not easy and as the activities required that he felt as though they were not simple. However, he considered that he thought and contacted his father more than usual during the study ("I did think of him more when I was doing the tasks") and reported that he felt closer to his father when doing the tasks and when receiving the materials from him. He also pointed out that this type of intervention would have been even more relevant if they were living further apart.

Karen considered that the activities made her think about her father in a different way and remember good moments from the past:

> *It made me think about my dad in a different way.* [Karen]

About receiving the media, she reported:

> *It was nice to see how much I was meaning to Dad (...) he was appreciating things. I felt loved and valued. You got a more intense relationship snapshot from us. It was interesting to see how we were interacting (...). I really enjoyed the way everything was handled. It made me feel closer to dad. It was something that brought me, my dad, and my brother together.* [Karen]

Study 2: Friendship Trial

Engagement and Patterns of Interaction

In phase 1, there was a high level of engagement with the media requests for all participants. Ronda and Irene responded to 100% (14/14) of the media requests, and Linda to 71% (10/14). Linda did not respond to any of the video requests because of difficulties using the video recording tool in the tablet, despite being trained to do so twice by the researchers. All participants interacted with the system on a daily basis, except for 1 request that Linda completed on the following day.

All participants preferred to interact with the system when they had time, rather than immediately after a request or parcel had arrived. Irene and Linda only responded to the requests immediately after the sound notification in 7% (1/14) of the times, and Ronda in 28% (4/14) of the times. On average, Ronda spent 4 min 23 seconds per interaction, Irene spent 3 min 51 seconds, and Linda spent 5 min 42 seconds. The replies that took the longest were those involving existing materials, such as pictures and objects.

In phase 2, the media sharing phase, the engagement of interaction with the app was 100% (9/9) for all participants. They interacted with the received media parcels on a daily basis. Regarding the sound reminder, Ronda and Irene interacted with the app in the same proportion before, after, or immediately after the reminder 33% (3/9) of the time. Linda interacted with the media parcels immediately after the sound alert 55% (5/9) of the time and before or after it 22% (2/9) of the time. This again reflects increased interest and anticipation in receiving media.

Type of Media and Media Content

As in study 1, the media parcels elicited from users were often more intimate than social media posts or even than things they might have said to each other face-to-face. In fact, some had the character of greeting cards that allowed people to express deeper feelings than what they are comfortable saying in person. For example, a request to record a message about her friendship with Ronda caused Irene to say the following:

> *She is a very kind friend, very caring, very dedicated. I think that regarding my friends I am privileged. Ronda is someone that I like very much.* [Irene]

When we asked Ronda to say how she felt about this to us, she wrote:

> *I appreciate the message very much and it made me emotional. I am what I am! I prioritize my friends because they are the shoulders that I lean on. Listening to the audio from Irene I feel safe and cared*

> *for. I know I can count on her friendship forever. Thank you, my friend.* [Ronda]

Regarding their favorite types of media, participants varied. While producing the media, Ronda liked pictures, writing, videos, and audios in that order. Her favorite tasks involved pictures because she retrieved them from old materials that reminded her of stories from the past that she had forgotten. One of her favorite tasks in the media collection phase was to talk about her favorites (singer, television show, and hobbies):

> *Because of that I started to listen to music that I loved but had forgotten.* [Ronda]

About the received media parcels, she reported that she did not have a favorite type and that the most important aspect was the content. She appreciated the messages she received from the friends:

> *I was moved and thought how true friendships are important in our lives. Who has friends has everything!* [Ronda]

She picked out 2 favorite media parcels: (1) the friends' childhood stories ("They reminded me of my own childhood. We usually do not share old stories like that"); and (2) their personal profiles. This was a composite parcel made by the researchers. The format was a magazine cover with each participant's photo and information that they had provided about themselves, in addition to the information provided from the friends to each other. She said that she really related to her own profile and was moved by it. She also thought that the profiles were a good snapshot of each of the friends.

Similar to Ronda, Irene also enjoyed producing the photos:

> *It was cool. I looked for old, forgotten photos from my friends but also family. It was good to revisit these materials.* [Irene]

Then her order of preference was audio, video, and text:

> *I do not like so much to show in pictures, so I did not like very much to show my face in the videos. But the audio was fine. I also don't like writing (...) I have never been good with the written word.* [Irene]

Her favorite activity was to record an audio about her childhood:

> *It made me remember my childhood, the people from my childhood (...) it was a good time of my life.* [Irene]

While receiving media, she said that she enjoyed everything and that the content was more important than the format. She noted that she especially enjoyed receiving the videos and pictures and that her favorite media parcel was an audio she received from Ronda:

> *I felt truly touched and even cried (...). I never expect compliments from other people and I don't even know if I am all that (...) I immediately picked up the phone and called her to say how much she means to me* [Irene]

For Linda, she enjoyed producing the types of media in the following order: text, photos, audio, and video. Although she likes to watch videos, she had difficulties using the device's camera to record videos by herself, which was frustrating for her. The media she most enjoyed producing was to find a snippet of a song to record:

> *I did not want to sing, but I enjoyed looking for a song that I liked to record a bit of it.* [Linda]

About receiving the media, she said that she appreciated all of them, irrespective of type.

Producing Media Versus Receiving Media

Feelings about preferring the producing or the receiving of the media parcels were different across participants. Irene preferred to receive the media:

> *It is more fun (...) you get to know more about the other person and their feelings (...) I felt emotional and even cried over some materials [messages from the friends].* [Irene]

Receiving the personal histories about the friends also made her realize that they had similar life experiences. Although she also enjoyed going through old photos because they reminded her of good times. Ronda reported that she enjoyed them both:

> *I loved them both (...) one phase completes the other (...). If you produce something and do not get anything back, you do not get closure. Producing the materials made me rethink my life, good times and difficult ones (...) and how I have overcome the difficulties (...). I visualized myself back in time, it was as if I was reliving it. I really enjoyed it. It is an interesting approach to carry out with my other friends as well.* [Ronda]

About the media sharing phase, she reported:

> *I always expected the messages as a surprise (...) the photos and audios moved me.* [Ronda]

As for Linda, she enjoyed producing the media parcels because she had the opportunity to revisit bits of her life and moments with the friends:

> *It touched me (...) I got in touch with my emotions, with personal stuff. It was a bit of work having to look for old materials, but it felt good to revive the friendship (...) it was getting cold* [Linda]

As in study 1, receiving the media in general produced more spontaneous and emotional responses about their feelings.

Feelings of Specific Social Connectedness

During the study, all participants reported feeling closer to each other, thinking about each other, and getting in touch more than usual. All participants commented that during the study, they were encouraged to talk about personal things and talk about how they felt about their friendship in a way that they would not naturally do. Irene said:

> *Sometimes it is hard for me to express certain feelings (...) this helped me to do that.* [Irene]

Ronda reported:

> *These talking points do not come to our conversations by chance. The sequence of questions brought memories (...) and the opportunity to express*

> *ourselves in a different way (...) the questions helped me to express my feelings.* [Ronda]

Although she had a stronger bond with Irene, she reported that she got more involved with Linda as well.

Irene reported that she felt closer to the friends in the media collection phase but even more in the media sharing phase. Irene was moved by the messages she received from both friends, and she decided to call them both:

> *I called Linda and I said, "I do not say it enough but I love you (...)." I also called Ronda to say that she is really important in my life and that I don't know what I would do without her.* [Irene]

After this call, Linda also called Irene once, and Irene called both friends to schedule a night out including 3 other friends. In this occasion, Linda spoke about a misunderstanding she had with Irene 10 years earlier that she still felt hurt about. Irene could not go to Linda's birthday, so she sent a gift through a friend. Linda did not appreciate it because she would rather have Irene come over in person to bring the present. Thus, she returned the gift and never talked about it again until this moment. They said that this was a good opportunity to set things straight and get closer to each other again, and it ended in a group hug. However, Irene felt a little hurt that Linda was still mad after such a long time. Ronda also commented that she and Irene will try to include Linda more in their social gatherings:

> *Because she has a boyfriend, we do not invite her a lot to go out with us (...) but she really enjoys our company and we will try to include her more.* [Ronda]

These reported feelings of closeness are reflected in a systematic rise in their RSDS scores, as shown in Multimedia Appendix 3, at the pretrial, midtrial, and posttrial periods. The only drop was between the mid- and posttrial periods for Irene's feelings of closeness with Linda. These went down because of the conflict over the birthday present reported above. This shows that developing relationships is a complex business. It may result in relationship work that raises contradictory feelings of closeness and distance within the same relationship pair.

Discussion

Principal Findings

We designed a novel communication system called *Media Parcels* that makes time-based requests for media and distributes them within a minimal network of just a few social contacts. The social networking system we designed aimed to support older people in deepening their social connections and relationships while still respecting their privacy concerns.

Broadly, our results indicated positive social effects for both deepening and developing of relationships. The elderly participants perceived personal and more intimate social connection as important to their lives, corroborating with other studies that present social integration with family, friends, or community as crucial to well-being in older age [6-9]. Our participants also appreciated the computer-mediated system as a means to encourage self-disclosure and initiate conversations about feelings toward one another, with more deep and meaningful content. This result supports the findings of Lindley et al [51] that explored the attitudes of older adults to keeping in touch with people who are important to them. The authors found that the older adults wanted to be in touch, and that staying connected was worthy of dedication. Moreover, the participants most valued the communication that is personally expressed, which requires a level of intensity that contrasts with the lightweight interaction that is increasingly afforded by new technologies. This was precisely fostered in our study by the *Media Parcels* system, in which social interaction topics were designed to encourage meaningful reflections and expressions of feelings. The importance of fostering relevant, meaningful social interactions is highlighted in studies by Carstensen et al [52,53], which suggest that older people are more motivated to derive emotional meaning from life and establish intimacy with other people, presenting a preference to invest in relationships that are emotionally rewarding and significant to them.

Testing the *Media Parcels* system on a trio of family members and in a separate trio of friends exposed values common to both networks as well as values exclusive to each social group. It also revealed new opportunities for computer-mediated communication.

Family Connections

The different family members showed differing levels of engagement with the system across media collection and delivery phases. The older person Paul was most enthusiastic overall, and his daughter Karen was a close second. However, his son Charles started out as a skeptic who found it difficult at first to respond to media requests about his father. All three became more enthusiastic in the second phase of the study when they received media parcels relating to Paul. Karen summed up their experience about the system by saying that it brought them all together in a way that would not have happened spontaneously. This is an important observation from the daughter, as the system was configured to focus on her father and his relationship with family following the death of his wife 6 months earlier. It shows that even bilateral media sharing of this kind stimulated reflections on family as a whole and discussion outside the system that affected all relationships. Thus, the system fostered a shift in the content of the communication between the older participant and his children, from basically discussing family affairs and obligations to talking about feelings toward one another and interests. This is particularly relevant as the complex nature of family ties, that includes feelings of both togetherness and responsibilities, may produce negative feelings with greater degrees of obligation and formality associated with familial relationships when compared with friendship [54]. Interestingly, participants told us they rarely discussed the media parcels themselves, in case they ruined the surprise of pending deliveries during the trial. The parcels were also about feelings that participants found hard to express or discuss in person.

Friendship Relations

Media sharing was made trilateral in the friendship trial, with three older people of differing closeness to each other. Again, we found various levels of engagement and different effects of the system across participants, but a positive endorsement from

all of them in the end. In fact, Ronda wrote to us after the trial saying how much she missed the rhythm of the parcel requests and deliveries and its new kind of connection to her friends. There was greater interest in media collection between friends than in the family trial, and all the participants perceived their friendship relations as precious and important to their well-being. In the specific case of the friendship between Ronda and Irene, they even spent time with each other more frequently than with family members and often offered emotional and health-related support to each other, because of either physical distance from close relatives or not desiring to be a source of concern to them. This points out the importance of having friendships later in life and is aligned to other studies that discuss the benefits of having peer relationships on the well-being of older people [55,56]. However, some points of contention were raised by the deliveries. This was most dramatically illustrated in the story of the rejected present between Irene and Linda, which raised a forgotten issue for Irene but resolved it for Linda. This example shows not only the power of the system and particular media requests to access facets of a relationship but also the dangers of doing that sometimes for relationship closeness.

Facilitation

A total of 2 types of facilitation seemed to be going on in the trials: (1) the media parcels themselves, as designed, were facilitating reflection and communication between parties and triggering further conversation outside the system; and (2) the human facilitator in the loop, with a background in clinical psychology, authored requests and selected responses to maximize positive effects on relationships.

Concerning the media parcels themselves, we found that participants could express feelings to each other that were hard to communicate face-to-face, as in the use of greetings cards. In this respect, media requests were effective relationship probes, revealing aspects of a relationship to other partners as well as to us. Finding and selecting media to share was also motivating, especially for Ronda and Linda in the friendship trial. This caused them to retrieve forgotten images and other materials and remember the good times in the light of intervening events (see also Frohlich et al [57]). These findings are similar to those in a recent study of media sharing to facilitate young people's conversations with relatives having dementia. The young people were encouraged to find media relating to the person with dementia through a system called *Ticket to Talk* and use these media as a kind of conversational playlist to stimulate conversation [58].

As for the human facilitator in the loop, human facilitation amplified the effect of media exchanges by personalizing them to the participants. It also continued outside the system, as the facilitator also conducted the interviews and became a sounding board for the participants' reflections on their relationships. In fact, the facilitator, the first author who is a trained clinical psychologist, realized that it could be a powerful tool for counseling. This raises issues for the future of the approach in terms of the levels of facilitation involved. If the intervention lasts longer than a few weeks, how personal should the media be to deepen relationships between particular people? And how important is it to have a professional facilitator designing and monitoring media?

Regarding the first issue, if a system such as *Media Parcels* should be used for longer periods, then the content of the media requests should be balanced with deep personal content and lightweight content, so that the media exchange does not become burdensome.

With respect to the second issue, in our study, the facilitator had a key role in designing the media requests and collecting and distributing the media produced by the participants. It is not uncommon to have human facilitators mediating technology or Web-based social interaction for vulnerable populations. For example, in Abrahão et al [35], the facilitators had a central role in helping the creation of digital storytelling by a care home elderly resident with dementia. Another example is the Scrapbook Circles network that is designed to allow disabled users to post content to friends and family through facilitators, if they wish. Similarly, the *Media Parcels* system could be easily scaled up to dozens of users by expanding the number of facilitators. However, personalized interventions such as the ones reported in this paper have a limit on scaling; a human facilitator dedicated to the task may be able to monitor and mediate from 3 to 4 dozens of groups weekly.

In contrast, if personalization is not a requirement, the system could be scaled up to hundreds and thousands of users using predetermined templates and algorithms, such as those used in recommender systems toward suggesting books and songs. Also, by indicating users' relationships in the system, it is feasible to create algorithms that automatize content distribution, so received media can be redirected without the need of facilitators. The ESPIM system used to deploy *Media Parcels* already allows the use of predefined templates and automatic media distribution. The figure of the human facilitator could also be removed by adding in the system functionality of direct communication between participants over a predefined set of conversational topics and media sharing requests. Future research could explore both options to reveal more about the role of media in deepening relationships at a distance, especially for the older population.

Communication

The communication dynamics of the *Media Parcels* system was unusual: (1) it was facilitated by a human who administered a time-based dialogue; (2) the pace of the facilitator-participant dialogue was slow, ranging from 1-3 times a day; (3) the explicit separation of the collection and delivery phases allowed us to assess the effects of each phase; (4) the distinct phases introduced delay into the communication, resulting in media exchange that was neither wholly asynchronous nor wholly synchronous; and (5) the introduction of a delay into the communication resulted in anticipation of media deliveries in which participants met the parcels as they arrived in real time.

Such communication dynamics could be adapted within the current system design. For example, the time-based dialogs could make use of different media requests between different configurations of people within a wider network, interleaved more closely in time with their delivery. Our findings from

study 1 support the evolution of an intervention approach that could be offered professionally to individuals as a kind of relationship therapy. Moreover, the findings from study 2 support a self-paced long-term approach of unfacilitated media sharing within small groups. To make this sustainable, message requests might come from friends themselves. In this case, the figure of the human facilitator could be removed by adding to the system features that include direct communication between participants, supported by algorithms processing an ever-growing context-based (eg, time and location) list of conversational topics and media sharing requests.

Conclusions

The Media Parcels system was an effective approach to promote media sharing of emotional content for the elderly population that participated in this research. Although the system was tested in only 2 trios of users, it was suitable to promote communication and deepen social relations between participants from different generations (as in study 1) or from the same generation (as in study 2). It can be expected that similar results could be generalizable to the elderly population with characteristics comparable with our sample.

In addition, the participants reported to feel motivated to produce and receive personal media content throughout the study. It is possible, though, that the interest in the use of the system could diminish over time because the system is not novel anymore or producing the media is too burdensome. Nevertheless, the social and personal nature of the media exchange could be motivating to keep users engaged in communication, especially if they feel socially isolated. Moreover, it could be easier for users to keep using the system for longer periods of time if the media requests for deep personal content are balanced with lightweight content and if the users are able to create their own threads of conversation. In fact, the use of a system such as Media Parcels is not necessarily intended to be of long-term. The length of use can be determined, for example, by a health care professional facilitator, focusing on shorter interactions between users according to therapeutic goals.

The Media Parcels design presents a novel solution for including older adults in social media sharing by introducing the concept of intimate directed continuous slow media sharing, which is different from the existing online communities. Our trial showed that the parcel metaphor applied to media content was easily understood by the older population, and the supporting computational system was easy enough to be quickly adopted.

As far as future study is concerned, the authors plan to conduct research expanding the number of participants in a group, targeting specifically older people that are classified as lonely or socially disconnected.

Acknowledgments

IZ was supported by São Paulo Research Foundation FAPESP (Grant Numbers 2017/09549-6 and 2015/18117-7). This study was also supported by funding from FAPESP (Grant Numbers 2016/00351-6 and 2016/50489-4), Conselho Nacional de Desenvolvimento Científico e Tecnológico (Grant Number 312058/2015-2), and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, Brazil—Finance Code001.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Relationship semantic differential scale.
[PDF File (Adobe PDF File), 39 KB - jmir_v21i10e14112_app1.pdf]

Multimedia Appendix 2
Participants' evaluation of their relationship toward one another according to RSDS scores.
[PNG File , 8 KB - jmir_v21i10e14112_app2.png]

Multimedia Appendix 3
Participants' evaluation of their relationship toward one another according to RSDS scores.
[PNG File , 20 KB - jmir_v21i10e14112_app3.png]

References

1. Lee RM, Robbins SB. The relationship between social connectedness and anxiety, self-esteem, and social identity. J Couns Psychol 1998;45(3):338-345 [FREE Full text] [doi: 10.1037/0022-0167.45.3.338]
2. Lee R, Steven B, Robbins SB. Understanding social connectedness in college women and men. J Couns Dev 2000;78(4):484-491. [doi: 10.1002/j.1556-6676.2000.tb01932.x]
3. van Bel DT, Smolders KC, IJsselsteijn WA, de Kort Y. Social connectedness: concept and measurement. Intel Environ 2009;2:67-74. [doi: 10.3233/978-1-60750-034-6-67]

4. Population Division - United Nations Population Division. 2017. World Population Prospects: Key Findings and Advance Tables URL: https://population.un.org/wpp/Publications/Files/WPP2017_KeyFindings.pdf [accessed 2019-08-15] [WebCite Cache ID 76i21AAN5]
5. World Health Organization. 2015. World Report on Ageing and Health URL: https://apps.who.int/iris/bitstream/handle/10665/186463/9789240694811_eng.pdf [accessed 2019-08-15] [WebCite Cache ID 76i2Jr0WS]
6. Bowling A, Dieppe P. What is successful ageing and who should define it? Br Med J 2005 Dec 24;331(7531):1548-1551 [FREE Full text] [doi: 10.1136/bmj.331.7531.1548] [Medline: 16373748]
7. Ward E, Barnes M, Gahagan B. The University of Brighton. 2012. Well-Being in Old Age: Findings From Participatory Research URL: https://research.brighton.ac.uk/en/publications/well-being-in-old-age-findings-from-participatory-research [accessed 2019-08-15] [WebCite Cache ID 76i1Iq7qe]
8. Grewal I, Lewis J, Flynn T, Brown J, Bond J, Coast J. Developing attributes for a generic quality of life measure for older people: preferences or capabilities? Soc Sci Med 2006 Apr;62(8):1891-1901. [doi: 10.1016/j.socscimed.2005.08.023] [Medline: 16168542]
9. McLaughlin SJ, Jette AM, Connell CM. An examination of healthy aging across a conceptual continuum: prevalence estimates, demographic patterns, and validity. J Gerontol A Biol Sci Med Sci 2012 Jun;67(7):783-789 [FREE Full text] [doi: 10.1093/gerona/glr234] [Medline: 22367432]
10. Cornwell B, Laumann EO, Schumm LP. The social connectedness of older adults: a national profile*. Am Sociol Rev 2008;73(2):185-203 [FREE Full text] [doi: 10.1177/000312240807300201] [Medline: 19018292]
11. World Health Organization. 2002. Active Ageing: A Policy Framework URL: https://apps.who.int/iris/bitstream/handle/10665/67215/WHO_NMH_NPH_02.8.pdf [accessed 2019-08-15] [WebCite Cache ID 76i2B1S9i]
12. Office for National Statistics. 2018. Internet Access – Households and Individuals, Great Britain: 2018 URL: https://www.ons.gov.uk/peoplepopulationandcommunity/householdcharacteristics/homeinternetandsocialmediausage/bulletins/internetaccesshouseholdsandindividuals/2018 [accessed 2019-08-15]
13. IBGE | Agência de Notícias. 2017. Continuous PNAD - ICT 2017: Internet Reaches Three of Every Four Households in the Country URL: https://agenciadenoticias.ibge.gov.br/en/agencia-press-room/2185-news-agency/releases-en/23453-continuous-pnad-ict-2017-internet-reaches-three-in-every-four-households-in-the-country [accessed 2019-08-15]
14. Pew Research Center. 2019. Mobile Fact Sheet URL: https://www.pewinternet.org/fact-sheet/mobile/ [accessed 2019-08-15]
15. Pimentel MG, Cunha BC, Antonelli HL, Rodrigues SS, Neto OJ, Rocha AC, et al. Enhancing Older Adults Connectivity by Introducing Mobile Devices Communication Tools. In: Proceedings of the 7th International Conference on Software Development and Technologies for Enhancing Accessibility and Fighting Info-Exclusion. 2016 Presented at: DSAI'16; December 1-3, 2016; Vila Real, Portugal p. 353-361 URL: https://dl.acm.org/citation.cfm?id=3019994 [doi: 10.1145/3019943.3019994]
16. Lindley SE, Harper R, Sellen A. Designing for Elders: Exploring the Complexity of Relationships in Later Life. In: Proceedings of the 22nd British HCI Group Annual Conference on People and Computers: Culture, Creativity, Interaction. 2008 Presented at: BCS-HCI'08; September 1-5, 2008; Liverpool, United Kingdom p. 77-86. [doi: 10.1145/1531514.1531525]
17. Pedell S, Vetere F, Kulik L, Ozanne E, Gruner A. Social Isolation of Older People: The Role of Domestic Technologies. In: Proceedings of the 22nd Conference of the Computer-Human Interaction Special Interest Group of Australia on Computer-Human Interaction. 2010 Presented at: OZCHI'10; November 22-26, 2010; Brisbane, Australia p. 164-167. [doi: 10.1145/1952222.1952255]
18. van Dijck J. Mediated Memories in the Digital Age. First Edition. Palo Alto, California: Stanford University Press; 2007.
19. Sarvas R, Frohlich DM. From Snapshots to Social Media - The Changing Picture of Domestic Photography. First Edition. London: Springer; 2011.
20. Dib L, Petrelli D, Whittaker S. Sonic Souvenirs: Exploring the Paradoxes of Recorded Sound for Family Remembering. In: Proceedings of the 2010 ACM Conference on Computer Supported Cooperative Work. 2010 Presented at: CSCW'10; February 6-10, 2010; Savannah, Georgia, USA p. 391-400. [doi: 10.1145/1718918.1718985]
21. Kirk DS, Sellen A. On human remains: values and practice in the home archiving of cherished objects. ACM Trans Comput-Hum Interact 2010 Jul 1;17(3):10-53. [doi: 10.1145/1806923.1806924]
22. Park N, Jin B, Jin SA. Effects of self-disclosure on relational intimacy in Facebook. Comput Human Behav 2011 Sep;27(5):1974-1983. [doi: 10.1016/j.chb.2011.05.004]
23. Orben AC, Dunbar RI. Social media and relationship development: the effect of valence and intimacy of posts. Comput Human Behav 2017 Aug;73:489-498. [doi: 10.1016/j.chb.2017.04.006]
24. Rebelo C. The use of the internet and Facebook by the elders in Portugal: an exploratory study. Observatorio 2015;9(3):129-153 [FREE Full text]
25. Chakraborty R, Vishik C, Rao HR. Privacy preserving actions of older adults on social media: exploring the behavior of opting out of information sharing. Decis Support Syst 2013 Nov;55(4):948-956. [doi: 10.1016/j.dss.2013.01.004]
26. Sinclair TJ, Grieve R. Facebook as a source of social connectedness in older adults. Comput Human Behav 2017 Jan;66(5):363-369. [doi: 10.1016/j.chb.2016.10.003]
27. Quinn K. Cognitive effects of social media use: a case of older adults. Soc Media Soc 2018 Jul 19;4(3):205630511878720. [doi: 10.1177/2056305118787203]

28. Chen YR, Schulz PJ. The effect of information communication technology interventions on reducing social isolation in the elderly: a systematic review. J Med Internet Res 2016 Jan 28;18(1):e18 [FREE Full text] [doi: 10.2196/jmir.4596] [Medline: 26822073]
29. Karimi A, Neustaedter C. From High Connectivity to Social Isolation: Communication Practices of Older Adults in the Digital Age. In: Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work Companion. 2012 Presented at: CSCW'12; February 11-15, 2012; Seattle, Washington, USA p. 127-130 URL: http://doi.acm.org/10.1145/2141512.2141559 [doi: 10.1145/2141512.2141559]
30. Taipale S, Farinosi M. The Big Meaning of Small Messages: The Use of WhatsApp in Intergenerational Family Communication. In: Proceedings of the International Conference on Human Aspects of IT for the Aged Population. 2018 Presented at: ITAP'18; July 15–20, 2018; Las Vegas, NV, USA p. 532-546 URL: https://link.springer.com/chapter/10.1007/978-3-319-92034-4_40 [doi: 10.1007/978-3-319-92034-4_40]
31. Waycott J, Davis H, Vetere F, Morgans A, Gruner A, Ozanne E, et al. Captioned Photographs in Psychosocial Aged Care: Relationship Building and Boundary Work. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2014 Presented at: CHI'14; April 26-May 1, 2014; Toronto, Ontario, Canada p. 4167-4176 URL: https://dl.acm.org/citation.cfm?id=2557290 [doi: 10.1145/2556288.2557290]
32. Cornejo R, Tentori M, Favela J. Enriching in-person encounters through social media: a study on family connectedness for the elderly. Int J Hum-Comput St 2013 Sep;71(9):889-899. [doi: 10.1016/j.ijhcs.2013.04.001]
33. Zargham S, Ćalić J, Frohlich DM. 4streams: An Ambient Photo Sharing Application for Extended Families. In: Proceedings of the 2015 British HCI Conference. 2015 Presented at: British HCI'15; July 13-17, 2015; Lincoln, Lincolnshire, United Kingdom p. 165-174 URL: https://dl.acm.org/citation.cfm?id=2783589 [doi: 10.1145/2783446.2783589]
34. Garattini C, Wherton J, Prendergast D. Linking the lonely: an exploration of a communication technology designed to support social interaction among older adults. Universal Access Inf Soc 2011 Jun 7;11(2):211-222. [doi: 10.1007/s10209-011-0235-y]
35. Abrahão AR, da Silva PF, Frohlich DM, Chrysanthaki T, Gratão A, Castro P. Mobile Digital Storytelling in a Brazilian Care Home. In: Proceedings of the International Conference on Human Aspects of IT for the Aged Population. 2018 Presented at: ITAP'18; July 15-20, 2018; Las Vegas, NV, USA p. 403-421. [doi: 10.1007/978-3-319-92034-4_31]
36. Zaine I, Rodrigues KR, Cunha BC, Viel CC, Orlando AF, Neto OJ, et al. ESPIM: An Ubiquitous Data Collection and Programmed Intervention System using ESM and Mobile Devices. In: Proceedings of the 22nd Brazilian Symposium on Multimedia and the Web. 2016 Presented at: Webmedia'16; November 8-11, 2016; Teresina, Piauí State, Brazil p. 13-14 URL: https://dl.acm.org/citation.cfm?id=2988222 [doi: 10.1145/2976796.2988222]
37. Viel CC, Rodrigues KR, Zaine I, Cunha BC, Scalco LF, Pimentel MG. Personalized Ubiquitous Data Collection and Intervention as Interactive Multimedia Documents. In: Proceedings of the 2017 ACM Symposium on Document Engineering. 2017 Presented at: DocEng'17; September 4-7, 2017; Valletta, Malta p. 223-226 URL: https://dl.acm.org/citation.cfm?id=3121046 [doi: 10.1145/3103010.3121046]
38. Csikszentmihalyi M, Larson R. Validity and reliability of the experience-sampling method. J Nerv Ment Dis 1987 Sep;175(9):526-536. [doi: 10.1097/00005053-198709000-00004] [Medline: 3655778]
39. Skinner BF. Programmed instruction revisited. Phi Delta Kappan 1986;68(2):103-110 [FREE Full text]
40. Cunha BC, Rodrigues KH, Zaine I, Scalco LF, Viel CC, Pimentel MG. Web-Based Authoring of Multimedia Intervention Programs for Mobile Devices: A Case Study on Elderly Digital Literacy. In: Proceedings of the 34th ACM/SIGAPP Symposium on Applied Computing. 2019 Presented at: SAC'19; April 8-12, 2019; Limassol, Cyprus p. 484-491. [doi: 10.1145/3297280.3297325]
41. Schlochtermeier LH, Pehrs C, Kappelhoff H, Kuchinke L, Jacobs AM. Emotion processing in different media types: realism, complexity, and immersion. J Syst Integr Neurosci 2015;1(2):41-47. [doi: 10.15761/JSIN.1000109]
42. Dougher MJ, Hayes SC. Clinical Behavior Analysis. Oakland, California: Context Press; 2000.
43. Gaver WW, Dunne A, Pacenti E. Design: cultural probes. Interactions 1999;6(1):21-29. [doi: 10.1145/291224.291235]
44. Folstein MF, Folstein SE, McHugh PR. 'Mini-mental state'. A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 1975 Nov;12(3):189-198. [doi: 10.1016/0022-3956(75)90026-6] [Medline: 1202204]
45. Osgood CE, Suci GJ, Tannenbaum P. The Measurement of Meaning. First Edition. Champaign, USA: University of Illinois Press; 1957.
46. Neves BB, Amaro F. Too old for technology? How the elderly of Lisbon use and perceive ICT. J Community Inform 2012;8(1):189-194 [FREE Full text]
47. Laugwitz B, Held T, Schrepp M. Construction and Evaluation of a User Experience Questionnaire. In: Proceedings of the Symposium of the Austrian HCI and Usability Engineering Group. 2008 Presented at: USAB'08; November 20-21, 2008; Graz, Austria p. 63-76. [doi: 10.1007/978-3-540-89350-9_6]
48. Brooke J. SUS: a quick and dirty usability scale. In: Jordan PW, Thomas B, Weerdmeester BA, McClelland IL, editors. Usability Evaluation In Industry. London: Taylor & Francis; 1996:189-194.
49. Díaz-Bossini JM, Moreno L. Accessibility to mobile interfaces for older people. Procedia Comput Sci 2014;27:57-66 [FREE Full text] [doi: 10.1016/j.procs.2014.02.008]

50. Petrovčič A, Taipale S, Rogelj A, Dolničar V. Design of mobile phones for older adults: an empirical analysis of design guidelines and checklists for feature phones and smartphones. Int J Hum-Comput Int 2017 Sep 5;34(3):251-264. [doi: 10.1080/10447318.2017.1345142]
51. Lindley SE, Harper R, Sellen A. Desiring to Be in Touch in a Changing Communications Landscape: Attitudes of Older Adults. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2009 Presented at: CHI'09; April 4-9, 2009; Boston, MA, USA p. 1493-1702. [doi: 10.1145/1518701.1518962]
52. Carstensen LL, Fung HH, Charles ST. Socioemotional selectivity theory and the regulation of emotion in the second half of life. Motiv Emot 2003;27(2):103-123 [FREE Full text] [doi: 10.1023/A:1024569803230]
53. Carstensen LL, Gross JJ, Fung HH. The social context of emotional experience. Annu Rev Gerontol Geriatr 1998;17(1):325-352 [FREE Full text] [doi: 10.1891/0198-8794.17.1.325]
54. Rook KS. Reciprocity of social exchange and social satisfaction among older women. J Pers Soc Psychol 1987;52(1):145-154 [FREE Full text] [doi: 10.1037/0022-3514.52.1.145]
55. Adams RG, Blieszner R. Aging well with friends and family. Am Behav Sci 2016 Jul 27;39(2):209-224 [FREE Full text] [doi: 10.1177/0002764295039002008]
56. Fiori KL, Antonucci TC, Cortina KS. Social network typologies and mental health among older adults. J Gerontol B Psychol Sci Soc Sci 2006 Jan 1;61(1):P25-P32. [doi: 10.1093/geronb/61.1.P25]
57. Frohlich DM, Wall S, Kiddle G. Rediscovery of forgotten images in domestic photo collections. Pers Ubiquitous Comput 2012 Nov 8;17(4):729-740 [FREE Full text] [doi: 10.1007/s00779-012-0612-4]
58. Welsh D, Morrissey K, Foley S, McNaney R, Salis C, McCarthy J, et al. Ticket to Talk: Supporting Conversation Between Young People and People with Dementia Through Digital Media. In: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. 2018 Presented at: CHI'18; April 21-26, 2018; Montreal QC, Canada. [doi: 10.1145/3173574.3173949]

Abbreviations

ESPIM: Experience Sampling and Programmed Intervention Method
RSDS: Relationship Semantic Differential Scale
SUS: System Usability Scale
UEQ: User Experience Questionnaire

Edited by G Eysenbach; submitted 29.03.19; peer-reviewed by J Favela, W Zhang, M Khazaee-Pool; comments to author 18.05.19; revised version received 12.07.19; accepted 21.07.19; published 03.10.19.

Please cite as:
Zaine I, Frohlich DM, Rodrigues KRDH, Cunha BCR, Orlando AF, Scalco LF, Pimentel MDGC
Promoting Social Connection and Deepening Relations Among Older Adults: Design and Qualitative Evaluation of Media Parcels
J Med Internet Res 2019;21(10):e14112
URL: https://www.jmir.org/2019/10/e14112
doi:10.2196/14112
PMID:31584001

Original Paper

Development of the Therapeutic Alliance and its Association With Internet-Based Mindfulness-Based Cognitive Therapy for Distressed Cancer Patients: Secondary Analysis of a Multicenter Randomized Controlled Trial

Else Bisseling[1*], MD; Linda Cillessen[1*], MSc; Philip Spinhoven[2,3*], PhD; Melanie Schellekens[4*], MSc, PhD; Félix Compen[1*], MSc; Marije van der Lee[4*], PhD; Anne Speckens[1*], MD, PhD

[1]Radboudumc for Mindfulness, Department of Psychiatry, Radboud University Medical Center, Nijmegen, Netherlands

[2]Institute of Psychology, Leiden University, Leiden, Netherlands

[3]Department of Psychiatry, Leiden University Medical Center, Leiden, Netherlands

[4]Helen Dowling Institute, Centre for Psycho-Oncology, Scientific Research Department, Bilthoven, Netherlands

[*]all authors contributed equally

Corresponding Author:
Else Bisseling, MD
Radboudumc for Mindfulness
Department of Psychiatry
Radboud University Medical Center
966/CvM, Postbus 9101
Nijmegen, 6500HB
Netherlands
Phone: 31 243615445
Email: else.bisseling@radboudumc.nl

Abstract

Background: Mindfulness-based cognitive therapy (MBCT) is an evidence-based group-based psychological treatment in oncology, resulting in reduction of depressive and anxiety symptoms. Internet-based MBCT (eMBCT) has been found to be an effective alternative for MBCT. The therapeutic alliance (the bond between therapist and patient,) is known to have a significant impact on psychological treatment outcomes, including MBCT. A primary concern in the practice of eMBCT is whether a good therapeutic alliance can develop. Although evidence for the beneficial effect of therapist assistance on treatment outcome in internet-based interventions (IBIs) is accumulating, it is still unclear whether the therapeutic alliance is related to outcome in IBIs.

Objective: This study aimed to (1) explore whether early therapeutic alliance predicts treatment dropout in MBCT or eMBCT, (2) compare the development of the therapeutic alliance during eMBCT and MBCT, and (3) examine whether early therapeutic alliance is a predictor of the reduction of psychological distress and the increase of mental well-being at posttreatment in both conditions.

Methods: This study was part of a multicenter randomized controlled trial (n=245) on the effectiveness of MBCT or eMBCT for distressed cancer patients. The therapeutic alliance was measured at the start of week 2 (ie, early therapeutic alliance), week 5, and week 9. Outcome measures were psychological distress, measured with the Hospital Anxiety and Depression Scale, and mental well-being, measured with the Mental Health Continuum-Short Form.

Results: The strength of early therapeutic alliance did not predict treatment dropout in MBCT or eMBCT (B=−.39; *P*=.21). Therapeutic alliance increased over time in both conditions ($F_{2,90}$=16.46; Wilks λ=0.732; *P*<.001). This increase did not differ between eMBCT and MBCT ($F_{1,91}$=0.114; *P*=.74). Therapeutic alliance at week 2 predicted a decrease in psychological distress (*B*=−.12; t_{114}=−2.656; *P*=.01) and an increase in mental well-being (*B*=.23; t_{113}=2.651; *P*=.01) at posttreatment. The relationship with reduction of psychological distress differed between treatments: a weaker early therapeutic alliance predicted higher psychological distress at posttreatment in MBCT but not in eMBCT (*B*=.22; t_{113}=2.261; *P*=.03).

Conclusions: A therapeutic alliance can develop in both eMBCT and MBCT. Findings revealed that the strength of early alliance did not predict treatment dropout. Furthermore, the level of therapeutic alliance predicted reduced psychological distress and increased mental well-being at posttreatment in both conditions. Interestingly, the strength of therapeutic alliance appeared to be more related to treatment outcome in group-based MBCT than in eMBCT.

Trial Registration: ClinicalTrials.gov NCT02138513; https://clinicaltrials.gov/ct2/show/NCT02138513

(J Med Internet Res 2019;21(10):e14065) doi:10.2196/14065

KEYWORDS

therapeutic alliance; telemedicine; mindfulness; cancer; patient dropouts

Introduction

Background

From 2025 onward, 20 million people worldwide will be diagnosed with cancer each year [1]. Approximately one-third of all cancer patients suffer from psychological distress [2], resulting in long-lasting reduced quality of life, decreased compliance with medical care, and prolonged duration of hospital stay [3,4]. Mindfulness-based interventions (MBIs) such as mindfulness-based cognitive therapy (MBCT) [5] are viable psychological treatment options for cancer patients. Mindfulness is defined as "Paying attention; on purpose, in the present moment and non-judgmentally" [6]. MBIs are effective in reducing psychological distress in cancer patients [7,8]. However, as MBIs typically require in-person attendance at a series of classes over several weeks, many cancer patients experience barriers to participate due to illness, side effects of anticancer treatment, fatigue, limited mobility, or transport options [9].

Internet-based MBCT (eMBCT), overcoming many of these barriers, was found to be as effective as face-to-face (f2f) group-based MBCT in reducing psychological distress and in improving positive mental health and quality of life in a multicenter randomized controlled trial (RCT) in cancer patients [10]. Although treatment dropout was higher in eMBCT than in MBCT, this did not influence treatment efficacy. In the long term, the reduction of psychological distress was significantly higher in eMBCT than in MBCT [11].

The effectiveness of a psychological intervention depends on therapist skills and the strength of the established therapeutic relationship or the therapeutic alliance [12,13]. Therapeutic alliance is defined as the collaborative and affective bond between the therapist and the patient [14,15]. A good alliance means that the patient and therapist are working well toward the goals of the therapy [16]. The quality of the therapeutic alliance can be attributed to the change in outcome of psychological interventions [17,18]. Alliance measures can assess this working relationship [16]. The therapeutic alliance is often assessed with the Working Alliance Inventory (WAI) [19], measuring the degree of mutual trust between the client and therapist, their agreement on treatment goals, and their agreement on how to reach these goals. Alliance is commonly measured early, middle, and late in therapy. Early alliance strength does seem to be associated with outcome [16]. The significant impact of therapeutic alliance on f2f psychotherapy outcomes has been demonstrated in several meta-analyses [18,20]. It was recently found that therapeutic alliance also predicts the outcome of MBCT in cancer patients [21].

A systematic review of the therapeutic alliance in internet-based interventions (IBIs) indicated the development of therapeutic alliance to be equivalent to its development in f2f therapy [22]. In addition, Reynolds found that clients (n=30) and therapists (n=30) rated their session impact and alliances in Web-based text psychotherapy as strong as previously found in f2f psychotherapy [23]. With regard to the association between therapeutic alliance and outcome, findings are mixed. Earlier research found an association between therapeutic alliance and outcome in internet-based cognitive behavioral therapy (CBT) for depression and symptoms of posttraumatic stress [24,25] and in blended CBT for depression [26]. In a sample of cancer survivors (n=46), it was found that a strong therapeutic alliance predicted depressive symptom reduction in telephone-assisted CBT [27]. However, in other studies, no significant relationship between treatment outcome and the therapeutic alliance was found [28,29]. As far as we know, no studies have been conducted on the association of therapeutic alliance and outcome in eMBCT, although online mindfulness interventions are increasingly used [30]. To date, findings in internet-based formats are uncertain, and the relationship between therapeutic alliance and outcome of internet-based psychological interventions remains understudied [31]. The main aim of this study was to contribute to the ongoing debate in psychotherapy research on the relative importance of the therapeutic relationship on treatment outcome in (internet-based) psychological interventions. Gaining knowledge about whether therapeutic aspects, such as therapeutic alliance, influence outcome in eMBCT might improve therapeutic strategies and adequate implementation strategies.

Objectives

The objectives of this study were to (1) explore whether early therapeutic alliance predicts treatment dropout in MBCT or eMBCT, (2) compare the development of the therapeutic alliance during eMBCT and MBCT, and (3) examine whether early therapeutic alliance is a predictor of the reduction of psychological distress and the increase of mental well-being at posttreatment in both conditions.

Methods

Design

This study was part of a large multicenter RCT (n=245) on the effectiveness of MBCT and eMBCT versus treatment as usual

(TAU) for distressed cancer patients. The study was registered on Clinicaltrials.gov (NCT02138513) shortly after the start of recruitment, reported following CONSORT guidelines [32], and a protocol paper was published in advance [10,33]. Participants were randomized to MBCT, eMBCT, or TAU. After 3 months, patients in TAU were randomly allocated to either MBCT or eMBCT.

This study involved data from all patients randomized to MBCT or eMBCT, including those patients who were randomized to MBCT or eMBCT after they had completed the 3-month TAU period. The local ethics committee approved this study (CMO Arnhem Nijmegen 2013/542). All participants provided written informed consent before study enrolment.

Study Population and Procedure

Patients were recruited in specialized mental health care institutes for psycho-oncology via social media, patient associations, and advertorials in local newspapers in the Netherlands. Patients who were interested in participation could enroll themselves at the study website for a screening assessment with the Hospital Anxiety and Depression Scale (HADS). Patients with a score of 11 or greater on the HADS were contacted by telephone by one of the researchers to assess eligibility. Inclusion criteria were as follows: (1) having any cancer diagnosis, (2) experiencing at least mild psychological distress, (3) computer literacy and access to the internet, (4) good command of the Dutch language, and (5) willingness to participate in either online or f2f group–based MBCT. Exclusion criteria were as follows: (1) severe psychiatric morbidity such as suicidal ideation or psychosis, (2) change in psychotropic medication within 3 months of baseline, and (3) current or previous participation in MBCT or mindfulness-based stress reduction. Before randomization, patients completed the baseline assessment.

Intervention and Therapists

MBIs, such as MBCT [5], are innovative and effective psychological treatment options and have increasingly been applied in somatic health care, including oncology. Mindfulness is defined as intentionally paying attention to present-moment experiences in an accepting and nonjudgmental way [34]. They were originally developed for patients with chronic somatic conditions to help them cope more effectively with their condition, by increasing the capacity to focus on present-moment experiences, even in the presence of internal discomfort, and furthermore to prevent relapse in patients with recurrent major depressive disorders [5]. More recently, it has been applied in other conditions [35]. A 2012 meta-analysis of 9 RCTs (n=955) of MBIs in cancer patients found significant improvements in depressive and anxiety symptoms [7]. Since then, a number of RCTs have replicated the reduction of distress following MBIs [36-40].

Group-Based Mindfulness-Based Cognitive Therapy

Patients randomized to group-based MBCT received the intervention according to the MBCT protocol of Segal et al [5]. The MBCT protocol was tailored to cancer patients by including cancer-related psychoeducation and adapted movement exercises. The group-based MBCT consisted of 8 weekly 2.5-hour group sessions, one 6-hour silent day, and daily home practice assignments guided by audio files. The sessions consisted of mindfulness practices, sharing experiences, and didactic teachings. Each participant received a folder with information on each session and a CD containing the audio files. The group-based MBCT was provided at the Radboud University Medical Center in Nijmegen, the Jeroen Bosch Hospital in 's-Hertogenbosch, and 4 mental health institutes specialized in psycho-oncology (Helen Dowling Institute in Bilthoven, Ingeborg Douwes Centrum in Amsterdam, de Vruchtenburg in Rotterdam/Leiden, and Het Behouden Huys in Haren).

Internet-Based Mindfulness-Based Cognitive Therapy

The eMBCT followed exactly the same protocol as the group-based MBCT. The eMBCT was mainly text based and included asynchronous written interaction with a therapist over email, similar to the study by Bruggeman-Everts et al [41]. Patients were given access to a website divided into a workspace containing 9 sessions (8 weeks+1 full-day silent retreat) and a secured inbox. The silent day consists of a 6-hour program of practicing mindfulness in silence, with various exercises, and a lunch and tea breaks in silence. The therapist initiated the eMBCT by sending a welcome message to the patient, personalized with a photo or brief description of oneself as introduction. When patients logged in, they were presented with the overview of all information and assignments due for that week. Each session contained an introductory text; daily formal (eg, sitting meditation) and informal exercises (eg, awareness of everyday activities); and other home practice, with guided audio files and accompanying diaries. They were shown (fictional) patients' descriptions to emphasize common experiences and clarify the use of the logfiles. The therapist provided written feedback on the completed logfiles via the secured inbox on a prearranged day of the week. Patients were given written instructions after week 5 to prepare for their silent day at home. In the week of the silent day, patients were provided with a program similar to the MBCT silent day. Having completed 4 or more sessions of MBCT was defined as a minimum adequate dose in both interventions [42].

Therapists

A total of 14 therapists participated: 7 taught both interventions, 2 taught only MBCT, and 5 taught only eMBCT. All therapists fulfilled the advanced criteria of the Association of Mindfulness-Based Teachers in the Netherlands and Flanders, which are in concordance with the UK Mindfulness-Based Teacher Trainer Network Good Practice Guidelines [43]. All therapists had previous experience in working with oncology patients and received a 2-day training workshop in the MBCT study protocol at the start of the study. Moreover, 2 additional day-long supervision meetings were organized during the intervention phase of 1 year and 6 months. Therapists without previous eMBCT experience were provided with guidelines and were supervised by more experienced eMBCT therapists.

Measures

Therapeutic Alliance

Therapeutic alliance was measured with a translated and shortened form of the *WAI* [19], which was administered at the start of week 2, week 5, and week 9. The WAI consists of 3 subscales assessing (1) how closely client and therapist agree on and are mutually engaged in the goals of treatment; (2) how closely client and therapist agree on how to reach the treatment goals; and (3) the degree of mutual trust, acceptance, and confidence between the client and therapist. Items were scored on a 5-point scale ranging from rarely (0) to always (5) [44,45]. The 12-item inventory was validated in a Dutch-speaking sample, showing an internal consistency of greater than 0.80 for all separate subscales and 0.87 for the total scale [46]. The scale was used before in oncology patients following eMBCT [47]. Internal consistency of the total scale of the version used in this study was high (Cronbach alpha=.93).

Psychological Distress

Psychological distress was measured with the 14-item *HADS* developed to measure depression and anxiety in general medical populations [32]. Items were scored on a 4-point scale ranging from rarely (0) to almost always (4). It has been validated in medical populations, including cancer patients [48], and has adequate psychometric properties to detect distress and to screen for psychiatric disorders in cancer patients [49,50]. Internal consistency of the total scale in the present sample was high (Cronbach alpha=.87). Patients completed the HADS before and after the intervention.

Mental Well-Being

Mental well-being was measured by the *Mental Health Continuum-Short Form* (MHC-SF) [51], which comprises 14 items, representing various feelings of well-being in the past month rated on a 6-point Likert scale from never (0), once or twice a week (1), about once a week (2), 2 or 3 times a week (3), almost every day (4), and every day (5). The MHC-SF contains 3 subscales: emotional, psychological, and social well-being. The short form of the MHC-SF has shown excellent internal consistency (>0.80). The test-retest reliability of the MHC-SF over 3 successive 3-month periods was 0.68, and the 9-month test-retest in a Dutch sample was 0.65 [52]. Internal consistency of the total scale in the sample used in this study was high (Cronbach alpha=.97). Patients completed the MHC-SF before and after the intervention.

Statistical Analyses

All analyses were conducted in SPSS version 25.0 (IBM Corp, 2017). Differences between conditions and between completers and noncompleters (<4 sessions attended) in demographic and clinical variables were tested with chi-square tests and *t* tests. To test whether strength of alliance predicted dropout of eMBCT and MBCT, binary logistic regression analyses were used.

For further analyses, we only included patients who completed MBCT or eMBCT (n=163). To test the development of the therapeutic alliance during eMBCT and MBCT, repeated-measures analyses of variance were used. To examine whether therapeutic alliance predicted outcome and whether this was different in the 2 conditions, separate hierarchical linear regression models were conducted. The dependent variable was posttreatment psychological distress (HADS) or mental well-being (MHC-SF). In the first step, independent variables were the baseline level of psychological distress or mental well-being and condition (MBCT/eMBCT). In the second step, therapeutic alliance at week 2 was included as predictor. In the third step, the interaction between condition and therapeutic alliance was added to test moderation.

Results

Participants

Of the cancer patients participating in the RCT (n=245), 125 were randomized to eMBCT and 120 were randomized to MBCT. Of all patients (n=125) randomized to eMBCT, 22 (19.2%) declined participation after randomization. Of all patients (n=120) randomized to MBCT, 25 (20.8%) declined participation after randomization (see Figure 1). Participants were mostly female, middle aged, and highly educated. Most patients were diagnosed with breast cancer and were treated with a curative intent (see Table 1).

Figure 1. Flowchart of the sample used in the present study. eMBCT: internet-based mindfulness-based cognitive therapy; MBCT: mindfulness-based cognitive therapy.

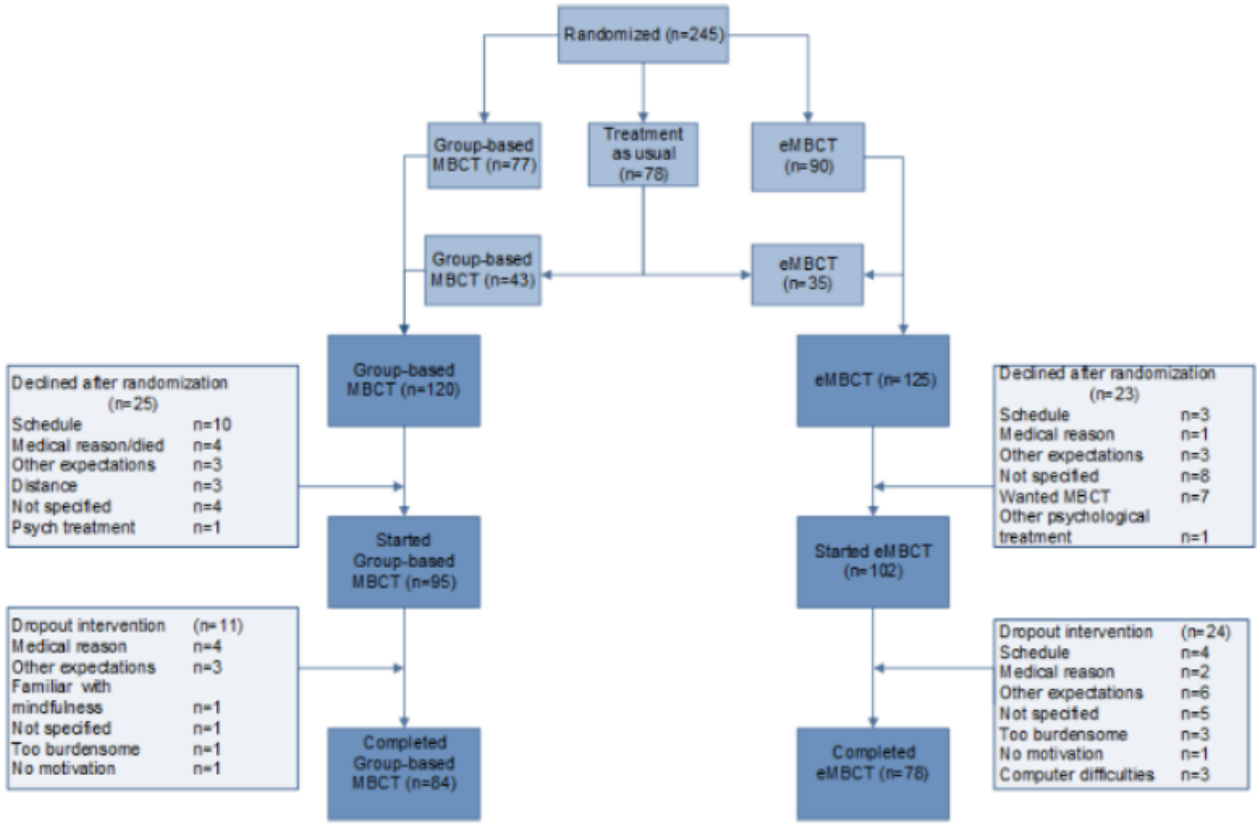


Table 1. Sample characteristics per intervention (mindfulness-based cognitive therapy/internet-based mindfulness-based cognitive therapy).

Characteristics	Mindfulness-based cognitive therapy[a] (n=120)	Internet-based mindfulness-based cognitive therapy[a] (n=125)	Chi-square (*df*)	*t* test (*df*)	*P* value
Sociodemographics					
Gender (female), n (%)	101 (84.2)	109 (87.2)	0.5 (1,244)	—[b]	.50
Relation (yes), n (%)	102 (85.0)	100 (80.0)	1.6 (1,244)	—	.40
Children, n (%)	79 (65.8)	90 (72.0)	1.1 (1,244)	—	.30
Education level, n (%)					
Low/middle	35 (29.2)	44 (35.4)	1.02 (1,244)	—	.31
High	85 (70.8)	81 (64.8)	1.02 (1,244)	—	.31
Age (years), mean (SD)	51.5 (11.1)	51.8 (10.2)	—	−0.18 (243)	.86
Clinical					
Cancer diagnosis, n (%)					
Breast	75 (62.5)	76 (60.8)	0.08 (1,244)	—	.78
Other	45 (37.5)	49 (39.2)	0.08 (1,244)	—	.78
Treatment intent, n (%)					
Curative	104 (86.7)	102 (81.6)	1.17 (1,244)	—	.28
Palliative	16 (13.3)	23 (18.4)	1.17 (1,244)	—	.28
Current anticancer treatment (yes), n (%)	56 (46.7)	60 (48.0)	0.44 (1,244)	—	.83
Psychiatric diagnosis (yes), n (%)	44 (36.0)	45 (36.7)	0.012 (1,244)	—	.91
Time since diagnosis, mean (SD)	3.5 (5.0)	3.4 (4.4)		0.21 (243)	.84
Psychological distress, Hospital Anxiety and Depression Scale, mean (SD)	18.2 (6.7)	16.9 (6.9)	1.49 (1,244)	—	.14
Mental well-being, Mental Health Continuum-Short Form, mean (SD)	35.0 (12.8)	37.4 (13.6)	−1.46 (1,244)	—	.15

[a]Both categories include the patients that were initially randomized to the treatment as usual group.

[b]Not applicable.

Dropout

Of the patients who started MBCT or eMBCT (n=198), 35 (17.7%) dropped out of the intervention. Dropout rate in eMBCT (n=24, 12.1%) was significantly higher than that in MBCT (n=11, 5.6%; *P*=.03). There were no differences in sociodemographics between participants who completed MBCT or eMBCT and those who did not.

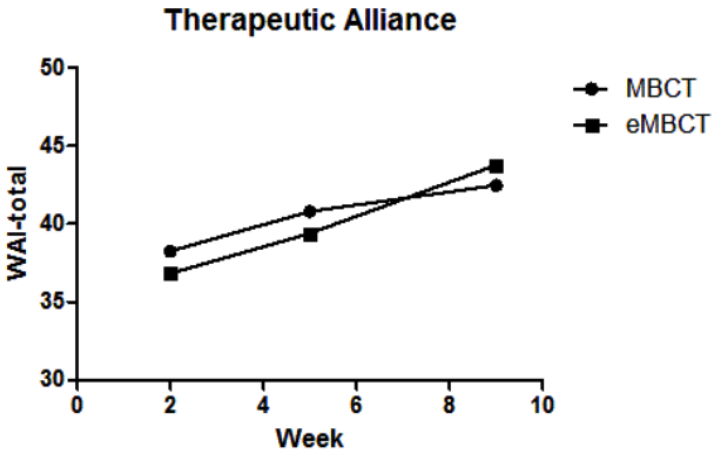
Patients who dropped out after the first sessions did not fill out the WAI questionnaire and, consequently, were not included in the analyses testing whether early therapeutic alliance predicted dropout of eMBCT and MBCT. However, reasons for dropout after session 1 were collected. Main reason was that the intervention was not as expected (n=6, 25.7%). Other reasons for dropout are summarized in Multimedia Appendix 1.

Analyses testing whether early therapeutic alliance predicted dropout of eMBCT and MBCT revealed that there was no significant difference in early therapeutic alliance between participants who dropped out of the intervention (mean 33.3, SD 7.9) and those who completed MBCT or eMBCT (mean 36.6, SD 9.5; *P*=.23). Moreover, this effect was not moderated by treatment condition, as there was no significant interaction between condition and the therapeutic alliance at the start of week 2 in association with dropout (*B*=−.13; *P*=.08).

Development of Therapeutic Alliance

As shown in Figure 2, level of therapeutic alliance increased significantly over time ($F_{2,90}$=16.46; Wilks =0.732; *P*<.001) and did not differ significantly between conditions (*P*=.74). There was no significant interaction between condition and time ($F_{2,90}$=1.636; Wilks =0.965; *P*=.20), implying that the development of therapeutic alliance did not differ between eMBCT and MBCT either.

Figure 2. Mean level of therapeutic alliance throughout mindfulness-based cognitive therapy and internet-based mindfulness-based cognitive therapy. eMBCT: internet-based mindfulness-based cognitive therapy; MBCT: mindfulness-based cognitive therapy.



Association of Therapeutic Alliance With Treatment Outcome

As you can see in Tables 2 and 3, early therapeutic alliance predicted both reduction of psychological distress (B=−.12; t114=−2.656; P=.01) and increase of mental well-being (B=.23; t113=2.651; P=.01) at posttreatment. Adding the interaction term between early therapeutic alliance and condition resulted in a significant increase in explained variance (ΔR^2=0.031; $F_{1,114}$=6.622; P=.01). The relationship between early therapeutic alliance and psychological distress at posttreatment was moderated by condition (B=.22; t_{113}=2.245; P=.03).

Examination of the interaction plot showed that participants of the group-based MBCT with a weaker early therapeutic alliance showed a significant higher level of psychological distress at posttreatment than those of eMBCT (P=.004; see also Multimedia Appendix 2).

Table 2. Relationship between therapeutic alliance and psychological distress: regression analyses. Italicized values indicate significance.

Variables	Regression analyses of psychological distress			Predictor		
	F value (*df*)	*P* value	Adjusted R^2	Beta	*t* test	*P* value
Psychological distress baseline	42.72 (2,115)	*<.001*	0.416	.58	8.69 (113)	*<.001*
Condition	42.72 (2,115)	*<.001*	0.416	−2.3	−2.36 (113)	*.01*
Therapeutic alliance week 2[a]	32.08 (3,114)	*<.001*	0.443	−.12	−2.57 (113)	*.01*
Therapeutic alliance week 2×condition[a]	26.17 (4,113)	*<.001*	0.463	.22	2.245 (113)	*.03*

[a]Corrected for baseline psychological distress and the condition, and the predictor×intervention interaction and psychological distress at posttreatment with separate hierarchical linear regressions (n=163).

Table 3. Relationship between therapeutic alliance and mental well-being: regression analyses. Italicized values indicate significance.

Variables	Regression analyses of mental well-being			Predictor		
	F value (df)	*P* value	Adjusted R^2	Beta	*t* test	*P* value
Mental well-being baseline	65.26 (2,115)	<.001	0.523	.76	10.84 (113)	*<.001*
Condition	65.26 (2,115)	<.001	0.523	1.9	1.16 (113)	.24
Therapeutic alliance week 2[a]	48.15 (3,114)	<.001	0.547	.24	2.66 (113)	*.01*
Therapeutic alliance week 2×condition[a]	36.07 (4,113)	<.001	0.545	−.12	−.697 (113)	.49

[a]Corrected for baseline mental well-being and the condition, and the predictor×intervention interaction and mental well-being at posttreatment with separate hierarchical linear regressions (n=163).

Discussion

Principal Findings

The aims of this study were to (1) explore whether the strength of early therapeutic alliance predicts treatment dropout in MBCT or eMBCT, (2) compare the development of the therapeutic alliance during eMBCT and MBCT, and (3) examine whether early therapeutic alliance is a predictor of the reduction of psychological distress and the increase of mental well-being at posttreatment in both conditions. Results showed that early alliance did not predict dropout and that the development of the

therapeutic alliance did not differ between eMBCT and MBCT. In addition, early therapeutic alliance predicted the reduction of psychological distress and improvement of mental well-being at posttreatment. Moreover, we found that the association between a weaker therapeutic alliance and smaller reduction of psychological distress was stronger in group-based MBCT than in eMBCT.

In our sample, dropout rates in eMBCT were somewhat higher than those in MBCT. This is in line with previous research findings that dropout rates in eMBCT are relatively high [10,53]. We could not confirm previous findings that a strong early alliance prevents patients from dropping out of psychological treatment [54]. In this respect, we would like to stress that our findings should be interpreted with caution. We were unable to include data from patients who dropped out after the first session because early therapeutic alliance was measured at the start of the second session. The main reason for dropping out after the first session was "other expectations." This finding stresses the importance of clearly explaining the treatment rationale and developing a shared understanding with patients at the start of MBCT or eMBCT. In our study, the therapist initiated the contact in eMBCT by sending a welcome message to the patient. As it was previously found in a qualitative study that in eMBCT asynchronicity of therapist contact/feedback is experienced by some patients as uncomfortable [9], adding an (f2f) appointment with the therapist at the start of eMBCT and synchronous therapist assistance in the early treatment phase might be useful to prevent dropout. Moreover, it was previously found that therapists seem concerned with the level of continuity of the training. Missing out on nonverbal information of patients made them unable to spot withdrawal at an early stage [9]. Alliance building may enhance engagement to treatment. So far, quantitative studies concerning the association of therapeutic alliance and adherence to IBIs are missing. Knowledge about alliance and adherence can be important to define the optimal level of therapist support in IBIs [31].

Regarding the association of therapeutic alliance and outcome, our findings are in line with previous (meta-analytic) research findings, showing that the therapeutic alliance is associated with outcome in f2f psychotherapy [18,20], including MBCT [21] and IBIs [24,25]. An explanation for the finding that the strength of the early alliance affects the outcome more in MBCT than in eMBCT could be that in a group-based setting, patients with a weak therapeutic alliance might conceal themselves in the group and might put in less effort to share experiences or practice. This might explain why psychological distress levels of those patients remained high compared with patients who experienced a strong therapeutic alliance. As a consequence, particularly in group-based settings, mindfulness therapists should be aware of their way of bonding with patients. Although we did not investigate this in our study, it has been previously found that for patients who have difficulties in building a relationship with others, which might be reflected in a weak therapeutic alliance, therapists who model effective ways of forming a strong alliance can process the formation of a strong (repaired) alliance [55,56], which in turn may lead to better outcome. Another explanation for this finding could be that in individual eMBCT, one's self-efficacy and intrinsic motivation are also important for treatment outcome than the experienced therapeutic alliance [11]. This is in line with previous qualitative findings in eMBCT for cancer patients, which revealed that participants stressed the importance of self-discipline and the individual nature as facilitators for eMBCT [9] suggesting that experienced motivation and discipline is also related to outcome.

Strengths and Limitations

A major strength of this study was that we directly compared the development of therapeutic alliance between patients participating in eMBCT and MBCT. Except for the delivery format, both interventions were similar in content. Moreover, we investigated the relationship between therapeutic alliance and treatment outcome both in terms of symptom reduction and mental well-being. Some limitations, however, need to be mentioned. At first, this study was conducted in a self-selected group of distressed cancer patients. Although this is in line with patients seeking help in clinical practice, this might limit the extent to which the findings of this study generalize to the broader population of cancer patients [57,58]. Second, an inclusion criterium was that participants would agree with randomization to an f2f or online intervention. Our inclusion criterium might have prevented patients that were not interested in online interventions from participating. This might limit the extent to which the present findings generalize to the broader population of cancer patients.

Research Implications

Further research should investigate the relation between therapeutic alliance and dropout more thoroughly, for instance, by measuring therapeutic alliance after each session. It would also be interesting to focus on examining what therapeutic techniques might be suitable to improve the development of a strong therapeutic alliance, both in eMBCT and MBCT. Further research in eMBCT could examine how therapeutic alliance can develop in different MBCT or eMBCT formats and how this influences treatment outcome. Treatment formats could vary from blended care to intensive synchronous individual therapist-assisted eMBCT to stand-alone eMBCT with computerized feedback.

Clinical Implications

As dropout rates were somewhat higher in eMBCT, efforts should be taken to prevent dropout. It is of interest to shed more light on how to encourage patients to adhere to the treatment. Building a good therapeutic alliance is of clinical importance in MBCT or eMBCT. By measuring therapeutic alliance as a routine outcome after every session, therapists can be encouraged to really tune in to patient's needs and adjust their therapy or their therapeutic attitude at an early treatment stage, which may benefit treatment outcome in both MBCT and eMBCT.

Acknowledgments

This research was funded by the Dutch national cancer foundation (KWF Kankerbestrijding; Grant no. 2012.WO14.C153). The authors would like to thank all patients and therapists who participated in this study and Heidi Willemse, Eva Witteveen, and David Huijts for their help in collecting data.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Dropout rates and reasons for dropout per condition.
[PDF File (Adobe PDF File), 93 KB - jmir_v21i10e14065_app1.pdf]

Multimedia Appendix 2
The association of early therapeutic alliance and outcome per condition.
[PNG File , 103 KB - jmir_v21i10e14065_app2.png]

References

1. Bray F. Transitions in human development and implications for cancer prevention and control. In: Stewart BW, Wild CP, editors. World Cancer Report 2014. Washington, DC: World Bank Publications; 2014:54-68.
2. Carlson L, Angen M, Cullum J, Goodey E, Koopmans J, Lamont L, et al. High levels of untreated distress and fatigue in cancer patients. Br J Cancer 2004 Jun 14;90(12):2297-2304 [FREE Full text] [doi: 10.1038/sj.bjc.6601887] [Medline: 15162149]
3. Mitchell AJ, Chan M, Bhatti H, Halton M, Grassi L, Johansen C, et al. Prevalence of depression, anxiety, and adjustment disorder in oncological, haematological, and palliative-care settings: a meta-analysis of 94 interview-based studies. Lancet Oncol 2011 Feb;12(2):160-174. [doi: 10.1016/S1470-2045(11)70002-X] [Medline: 21251875]
4. Mitchell AJ, Ferguson DW, Gill J, Paul J, Symonds P. Depression and anxiety in long-term cancer survivors compared with spouses and healthy controls: a systematic review and meta-analysis. Lancet Oncol 2013 Jul;14(8):721-732. [doi: 10.1016/S1470-2045(13)70244-4]
5. Segal ZV, Williams JM, Teasdale JD. Mindfulness-Based Cognitive Therapy for Depression: A New Approach to Preventing Relapse. New York: Guildford Press; 2013.
6. Kabat-Zinn J. Full Catastrophe Living: How to Cope With Stress, Pain and Illness Using Mindfulness Meditation. London: Piatkus Books; 2013.
7. Piet J, Würtzen H, Zachariae R. The effect of mindfulness-based therapy on symptoms of anxiety and depression in adult cancer patients and survivors: a systematic review and meta-analysis. J Consult Clin Psychol 2012 Dec;80(6):1007-1020. [doi: 10.1037/a0028329] [Medline: 22563637]
8. Zainal NZ, Booth S, Huppert FA. The efficacy of mindfulness-based stress reduction on mental health of breast cancer patients: a meta-analysis. Psychooncology 2013 Jul;22(7):1457-1465. [doi: 10.1002/pon.3171] [Medline: 22961994]
9. Compen FR, Bisseling EM, Schellekens MP, Jansen ET, van der Lee ML, Speckens AE. Mindfulness-based cognitive therapy for cancer patients delivered via internet: qualitative study of patient and therapist barriers and facilitators. J Med Internet Res 2017 Dec 18;19(12):e407 [FREE Full text] [doi: 10.2196/jmir.7783] [Medline: 29254912]
10. Compen F, Bisseling EM, Schellekens MP, Donders R, Carlson L, van der Lee M, et al. Face-to-face and internet-based mindfulness-based cognitive therapy compared with treatment as usual in reducing psychological distress in patients with cancer: a multicenter randomized controlled trial. J Clin Oncol 2018 Aug 10;36(23):2413-2421. [doi: 10.1200/JCO.2017.76.5669] [Medline: 29953304]
11. Cillessen L, Schellekens MP, van de Ven MO, Donders AR, Compen FR, Bisseling EM, et al. Consolidation and prediction of long-term treatment effect of group and online mindfulness-based cognitive therapy for distressed cancer patients. Acta Oncol 2018 Oct;57(10):1293-1302. [doi: 10.1080/0284186X.2018.1479071] [Medline: 29932784]
12. Kissane DW, Bloch S, Smith GC, Miach P, Clarke DM, Ikin J, et al. Cognitive-existential group psychotherapy for women with primary breast cancer: a randomised controlled trial. Psychooncology 2003 Sep;12(6):532-546. [doi: 10.1002/pon.683] [Medline: 12923794]
13. Wise TN, Biondi M, Costantini A. Psycho-Oncology. Washington, DC: American Psychiatric Publishing; 2013.
14. Bordin ES. Theory and research on the therapeutic working alliance: new directions. In: Horvath AO, Greenberg LS, editors. The Working Alliance: Theory, Research, and Practice. Hoboken, New Jersey, United States: Wiley; 1994:13-37.
15. Luborsky L. Therapeutic alliances as predictors of psychotherapy outcomes: factors explaining the predictive success. In: Horvath AO, Greenberg LS, editors. The Working Alliance: Theory, Research, and Practice. Hoboken, New Jersey, United States: Wiley; 1994:38-50.

16. Muran CJ, Barber JP, editors. The Therapeutic Alliance: An Evidence-Based Guide to Practice. New York, United States: The Guilford Press; 2010.
17. Horvath AO, Symonds BD. Relation between working alliance and outcome in psychotherapy: a meta-analysis. J Couns Psychol 1991;38(2):139-149. [doi: 10.1037/0022-0167.38.2.139]
18. Horvath AO, del Re AC, Flückiger C, Symonds D. Alliance in individual psychotherapy. Psychotherapy (Chic) 2011 Mar;48(1):9-16. [doi: 10.1037/a0022186] [Medline: 21401269]
19. Horvath AO, Greenberg LS. Development and validation of the working alliance inventory. J Couns Psychol 1989 Apr;36(2):223-233. [doi: 10.1037/0022-0167.36.2.223]
20. Martin DJ, Garske JP, Davis MK. Relation of the therapeutic alliance with outcome and other variables: a meta-analytic review. J Consult Clin Psychol 2000 Jun;68(3):438-450. [doi: 10.1037/0022-006X.68.3.438] [Medline: 10883561]
21. Bisseling EM, Schellekens MP, Spinhoven P, Compen FR, Speckens AE, van der Lee ML. Therapeutic alliance-not therapist competence or group cohesion-contributes to reduction of psychological distress in group-based mindfulness-based cognitive therapy for cancer patients. Clin Psychol Psychother 2019 May;26(3):309-318 [FREE Full text] [doi: 10.1002/cpp.2352] [Medline: 30650245]
22. Sucala M, Schnur JB, Constantino MJ, Miller SJ, Brackman EH, Montgomery GH. The therapeutic relationship in e-therapy for mental health: a systematic review. J Med Internet Res 2012 Aug 2;14(4):e110 [FREE Full text] [doi: 10.2196/jmir.2084] [Medline: 22858538]
23. Reynolds Jr DJ, Stiles WB, Bailer AJ, Hughes MR. Impact of exchanges and client-therapist alliance in online-text psychotherapy. Cyberpsychol Behav Soc Netw 2013 May;16(5):370-377 [FREE Full text] [doi: 10.1089/cyber.2012.0195] [Medline: 23530546]
24. Nordgren LB, Carlbring P, Linna E, Andersson G. Role of the working alliance on treatment outcome in tailored internet-based cognitive behavioural therapy for anxiety disorders: randomized controlled pilot trial. JMIR Res Protoc 2013 Jan 18;2(1):e4 [FREE Full text] [doi: 10.2196/resprot.2292] [Medline: 23612437]
25. Wagner B, Brand J, Schulz W, Knaevelsrud C. Online working alliance predicts treatment outcome for posttraumatic stress symptoms in Arab war-traumatized patients. Depress Anxiety 2012 Jul;29(7):646-651. [doi: 10.1002/da.21962] [Medline: 22678971]
26. Ly KH, Topooco N, Cederlund H, Wallin A, Bergström J, Molander O, et al. Smartphone-supported versus full behavioural activation for depression: a randomised controlled trial. PLoS One 2015;10(5):e0126559 [FREE Full text] [doi: 10.1371/journal.pone.0126559] [Medline: 26010890]
27. Applebaum AJ, DuHamel KN, Winkel G, Rini C, Greene PB, Mosher CE, et al. Therapeutic alliance in telephone-administered cognitive-behavioral therapy for hematopoietic stem cell transplant survivors. J Consult Clin Psychol 2012 Oct;80(5):811-816 [FREE Full text] [doi: 10.1037/a0027956] [Medline: 22468908]
28. Knaevelsrud C, Maercker A. Does the quality of the working alliance predict treatment outcome in online psychotherapy for traumatized patients? J Med Internet Res 2006 Dec 19;8(4):e31 [FREE Full text] [doi: 10.2196/jmir.8.4.e31] [Medline: 17213049]
29. Preschl B, Maercker A, Wagner B. The working alliance in a randomized controlled trial comparing online with face-to-face cognitive-behavioral therapy for depression. BMC Psychiatry 2011 Dec 6;11(1):49-57. [doi: 10.1186/1471-244x-11-189]
30. Plaza I, Demarzo MM, Herrera-Mercadal P, García-Campayo J. Mindfulness-based mobile applications: literature review and analysis of current features. JMIR Mhealth Uhealth 2013 Nov 1;1(2):e24 [FREE Full text] [doi: 10.2196/mhealth.2733] [Medline: 25099314]
31. Pihlaja S, Stenberg J, Joutsenniemi K, Mehik H, Ritola V, Joffe G. Therapeutic alliance in guided internet therapy programs for depression and anxiety disorders - a systematic review. Internet Interv 2018 Mar;11:1-10 [FREE Full text] [doi: 10.1016/j.invent.2017.11.005] [Medline: 30135754]
32. Schulz KF, Altman DG, Moher D, CONSORT Group. Consort 2010 Statement: updated guidelines for reporting parallel group randomised trials. BMC Med 2010 Mar 24;8:18 [FREE Full text] [doi: 10.1186/1741-7015-8-18] [Medline: 20334633]
33. Compen F, Bisseling EM, van der Lee ML, Adang EM, Donders AR, Speckens AE. Study protocol of a multicenter randomized controlled trial comparing the effectiveness of group and individual internet-based mindfulness-based cognitive therapy with treatment as usual in reducing psychological distress in cancer patients: the BeMind study. BMC Psychol 2015;3:27 [FREE Full text] [doi: 10.1186/s40359-015-0084-1] [Medline: 26273472]
34. Kabat-Zinn J. Full Catastrophe Living: Using the Wisdom of Your Body and Mind to Face Stress, Pain, and Illness. New York, United States: Bantam Publishing; 1990.
35. Crane R. Mindfulness-Based Cognitive Therapy: Distinctive Features. Abingdon, United Kingdom: Routledge; 2017.
36. Carlson LE, Doll R, Stephen J, Faris P, Tamagawa R, Drysdale E, et al. Randomized controlled trial of mindfulness-based cancer recovery versus supportive expressive group therapy for distressed survivors of breast cancer. J Clin Oncol 2013 Sep 1;31(25):3119-3126. [doi: 10.1200/JCO.2012.47.5210] [Medline: 23918953]
37. Garland SN, Carlson LE, Stephens AJ, Antle MC, Samuels C, Campbell TS. Mindfulness-based stress reduction compared with cognitive behavioral therapy for the treatment of insomnia comorbid with cancer: a randomized, partially blinded, noninferiority trial. J Clin Oncol 2014 Feb 10;32(5):449-457. [doi: 10.1200/JCO.2012.47.7265] [Medline: 24395850]

38. Lengacher CA, Reich RR, Paterson CL, Ramesar S, Park JY, Alinat C, et al. Examination of broad symptom improvement resulting from mindfulness-based stress reduction in breast cancer survivors: a randomized controlled trial. J Clin Oncol 2016 Aug 20;34(24):2827-2834 [FREE Full text] [doi: 10.1200/JCO.2015.65.7874] [Medline: 27247219]
39. Johannsen M, O'Connor M, O'Toole MS, Jensen AB, Højris I, Zachariae R. Efficacy of mindfulness-based cognitive therapy on late post-treatment pain in women treated for primary breast cancer: a randomized controlled trial. J Clin Oncol 2016 Oct 1;34(28):3390-3399. [doi: 10.1200/JCO.2015.65.0770] [Medline: 27325850]
40. Schellekens MP, van den Hurk DG, Prins JB, Donders AR, Molema J, Dekhuijzen R, et al. Mindfulness-based stress reduction added to care as usual for lung cancer patients and/or their partners: a multicentre randomized controlled trial. Psychooncology 2017 Dec;26(12):2118-2126. [doi: 10.1002/pon.4430] [Medline: 28337821]
41. Bruggeman Everts FZ, van der Lee ML, de Jager ME. Web-based individual mindfulness-based cognitive therapy for cancer-related fatigue — a pilot study. Internet Interv 2015 May;2(2):200-213 [FREE Full text] [doi: 10.1016/j.invent.2015.03.004]
42. Crane C, Williams JM. Factors associated with attrition from mindfulness-based cognitive therapy in patients with a history of suicidal depression. Mindfulness (N Y) 2010 Mar;1(1):10-20 [FREE Full text] [doi: 10.1007/s12671-010-0003-8] [Medline: 21125023]
43. Crane RS, Kuyken W, Williams JM, Hastings RP, Cooper L, Fennell MJ. Competence in teaching mindfulness-based courses: concepts, development and assessment. Mindfulness (N Y) 2012 Mar;3(1):76-84 [FREE Full text] [doi: 10.1007/s12671-011-0073-2] [Medline: 23293683]
44. Stinckens N, Ulburghs A, Claes L. HAN University of Applied Sciences. 2009. De Werkalliantie Als Sleutelelement in Het Therapiegebeuren: Meting Met Behulp Van De Wav-12, De Nederlandstalige Verkorte Versie Van De Working Alliance Inventory URL: https://www.han.nl/onderzoek/werkveld/projecten/alliantiegericht-werken-v/werkwijze/_attachments/2009_-_stinckens_ea_-_de_werkalliantie_als_sleuteleleement_in_het_therapiegebeuren_1_.pdf [accessed 2019-08-27]
45. Hatcher RL, Gillaspy JA. Development and validation of a revised short version of the working alliance inventory. Psychother Res 2006 Jan 1;16(1):12-25. [doi: 10.1080/10503300500352500]
46. Janse P, Boezen-Hilberdink L, van Dijk MK, Verbraak MJ, Hutschemaekers GJ. Measuring feedback from clients. Eur J Psychol Assess 2014 Jan;30(2):86-92. [doi: 10.1027/1015-5759/a000172]
47. Wolvers MD, Bruggeman-Everts FZ, van der Lee ML, van de Schoot R, Vollenbroek-Hutten MM. Effectiveness, mediators, and effect predictors of internet interventions for chronic cancer-related fatigue: the design and an analysis plan of a 3-armed randomized controlled trial. JMIR Res Protoc 2015 Jun 23;4(2):e77 [FREE Full text] [doi: 10.2196/resprot.4363] [Medline: 26104114]
48. Bjelland I, Dahl A, Haug T, Neckelmann D. The validity of the hospital anxiety and depression scale. An updated literature review. J Psychosom Res 2002 Feb;52(2):69-77. [doi: 10.1016/S0022-3999(01)00296-3] [Medline: 11832252]
49. Mitchell AJ, Meader N, Symonds P. Diagnostic validity of the hospital anxiety and depression scale (HADS) in cancer and palliative settings: a meta-analysis. J Affect Disord 2010 Nov;126(3):335-348. [doi: 10.1016/j.jad.2010.01.067] [Medline: 20207007]
50. Vodermaier A, Linden W, Siu C. Screening for emotional distress in cancer patients: a systematic review of assessment instruments. J Natl Cancer Inst 2009 Nov 4;101(21):1464-1488 [FREE Full text] [doi: 10.1093/jnci/djp336] [Medline: 19826136]
51. Keyes CL. Mental illness and/or mental health? Investigating axioms of the complete state model of health. J Consult Clin Psychol 2005 Jun;73(3):539-548. [doi: 10.1037/0022-006X.73.3.539] [Medline: 15982151]
52. Lamers SM, Westerhof G, Bohlmeijer ET, ten Klooster PM, Keyes CL. Evaluating the psychometric properties of the mental health continuum-short form (MHC-SF). J Clin Psychol 2011 Jan;67(1):99-110. [doi: 10.1002/jclp.20741] [Medline: 20973032]
53. Bruggeman-Everts FZ, Wolvers MD, van de Schoot R, Vollenbroek-Hutten MM, van der Lee ML. Effectiveness of two web-based interventions for chronic cancer-related fatigue compared to an active control condition: results of the 'fitter na kanker' randomized controlled trial. J Med Internet Res 2017 Oct 19;19(10):e336 [FREE Full text] [doi: 10.2196/jmir.7180] [Medline: 29051138]
54. Swift JK, Greenberg RP. Premature discontinuation in adult psychotherapy: a meta-analysis. J Consult Clin Psychol 2012 Aug;80(4):547-559. [doi: 10.1037/a0028226] [Medline: 22506792]
55. Tsai M, Kohlenberg RJ, Kanter JW. A functional analytic psychotherapy (FAP) approach to the therapeutic alliance. In: Muran JC, Barber JP, editors. The Therapeutic Alliance: An Evidence-Based Guide to Practice. Washington, DC: The Guilford Press; 2010:172-190.
56. Zilcha-Mano S, Errázuriz P. Early development of mechanisms of change as a predictor of subsequent change and treatment outcome: the case of working alliance. J Consult Clin Psychol 2017 May;85(5):508-520. [doi: 10.1037/ccp0000192] [Medline: 28345940]
57. Berger T. The therapeutic alliance in internet interventions: a narrative review and suggestions for future research. Psychother Res 2017 Sep;27(5):511-524. [doi: 10.1080/10503307.2015.1119908] [Medline: 26732852]

58. Berger T, Boettcher J, Caspar F. Internet-based guided self-help for several anxiety disorders: a randomized controlled trial comparing a tailored with a standardized disorder-specific approach. Psychotherapy (Chic) 2014 Jun;51(2):207-219. [doi: 10.1037/a0032527] [Medline: 24041199]

Abbreviations

CBT: cognitive behavioral therapy
eMBCT: internet-based mindfulness-based cognitive therapy
f2f: face-to-face
HADS: Hospital Anxiety and Depression Scale
IBI: internet-based intervention
MBCT: mindfulness-based cognitive therapy
MBI: mindfulness-based intervention
RCT: randomized controlled trial
TAU: treatment as usual
WAI: Working Alliance Inventory

Edited by P Bamidis; submitted 19.03.19; peer-reviewed by N Papaxristou, S Reppou; comments to author 25.04.19; revised version received 18.06.19; accepted 03.07.19; published 18.10.19.

Please cite as:
Bisseling E, Cillessen L, Spinhoven P, Schellekens M, Compen F, van der Lee M, Speckens A
Development of the Therapeutic Alliance and its Association With Internet-Based Mindfulness-Based Cognitive Therapy for Distressed Cancer Patients: Secondary Analysis of a Multicenter Randomized Controlled Trial
J Med Internet Res 2019;21(10):e14065
URL: https://www.jmir.org/2019/10/e14065
doi:10.2196/14065
PMID:31628791

Original Paper

Internet-Based Cognitive Behavioral Therapy for Chronic Fatigue Syndrome Integrated in Routine Clinical Care: Implementation Study

Margreet Worm-Smeitink[1,2,3], MSc; Arno van Dam[4,5], PhD; Saskia van Es[6], PhD; Rosalie van der Vaart[7], PhD; Andrea Evers[7,8], Prof Dr; Michel Wensing[9,10], Prof Dr; Hans Knoop[1,11], Prof Dr

[1]Expert Center for Chronic Fatigue, Department of Medical Psychology, University Medical Centers, Vrije Universiteit, Amsterdam, Netherlands
[2]Department of Medical Psychology, Radboud University Medical Center, Nijmegen, Netherlands
[3]Specialist Center for Complex Medically Unexplained Symptoms and Somatic Symptom Disorders, Dimence, Deventer, Netherlands
[4]Tranzo, School of Social and Behavioural Sciences, Tilburg University, Tilburg, Netherlands
[5]GGZ-Westelijk Noord Brabant, Institute for Mental Health, Bergen op Zoom, Netherlands
[6]PsyQ Somatiek en Psyche, Parnassia Groep, Amsterdam, Netherlands
[7]Health, Medical and Neuropsychology Unit, Institute of Psychology, Leiden University, Leiden, Netherlands
[8]Department of Psychiatry, Leiden University Medical Center, Leiden, Netherlands
[9]Department of General Practice and Health Services Research, Heidelberg University Hospital, Heidelberg, Germany
[10]Radboud Institute of Health Sciences, Radboud University Medical Center, Nijmegen, Netherlands
[11]Department of Medical Psychology, Amsterdam University Medical Centers, University of Amsterdam, Amsterdam, Netherlands

Corresponding Author:
Hans Knoop, Prof Dr
Department of Medical Psychology
Amsterdam University Medical Centers
University of Amsterdam
PO Box 22660, 1100DD
Amsterdam,
Netherlands
Phone: 31 20 4443925
Email: hans.knoop@amsterdamumc.nl

Abstract

Background: In a clinical trial, internet-based cognitive behavioral therapy (I-CBT) embedded in stepped care was established as noninferior to face-to-face cognitive behavioral therapy (CBT) for chronic fatigue syndrome (CFS). However, treatment effects observed in clinical trials may not necessarily be retained after implementation.

Objective: This study aimed to investigate whether stepped care for CFS starting with I-CBT, followed by face-to-face CBT, if needed, was also effective in routine clinical care. Another objective was to explore the role of therapists' attitudes toward electronic health (eHealth) and manualized treatment on treatment outcome.

Methods: I-CBT was implemented in five mental health care centers (MHCs) with nine treatment sites throughout the Netherlands. All patients with CFS were offered I-CBT, followed by face-to-face CBT if still severely fatigued or disabled after I-CBT. Outcomes were the Checklist Individual Strength, physical and social functioning (Short-Form 36), and limitations in daily functioning according to the Work and Social Adjustment Scale. The change scores (pre to post stepped care) were compared with a benchmark: stepped care from a randomized controlled trial (RCT) testing this treatment format. We calculated correlations of therapists' attitudes toward manualized treatment and eHealth with reduction of fatigue severity.

Results: Overall, 100 CFS patients were referred to the centers. Of them, 79 started with I-CBT, 20 commenced directly with face-to-face CBT, and one did not start at all. After I-CBT, 48 patients met step-up criteria; of them, 11 stepped up to face-to-face CBT. Increase in physical functioning (score of 13.4), social functioning (20.4), and reduction of limitations (10.3) after stepped care delivered in routine clinical care fell within the benchmarks of the RCT (95% CIs: 12.8-17.6; 25.2-7.8; and 7.4-9.8, respectively). Reduction of fatigue severity in the MHCs was smaller (12.6) than in the RCT (95% CI 13.2-16.5). After I-CBT only, reduction of fatigue severity (13.2) fell within the benchmark of I-CBT alone (95% CI 11.1-14.2). Twenty therapists treated

between one and 18 patients. Therapists were divided into two groups: one with the largest median reduction of fatigue and one with the smallest. Patients treated by the first group had a significantly larger reduction of fatigue severity (15.7 vs 9.0; t=2.42; P=.02). There were no (statistically significant) correlations between therapists' attitudes and reduction in fatigue.

Conclusions: This study is one of the first to evaluate stepped care with I-CBT as a first step in routine clinical care. Although fatigue severity and disabilities were reduced, reduction of fatigue severity appeared smaller than in the clinical trial. Further development of the treatment should aim at avoiding dropout and encouraging stepping up after I-CBT with limited results. Median reduction of fatigue severity varied largely between therapists. Further research will help understand the role of therapists' attitudes in treatment outcome.

(J Med Internet Res 2019;21(10):e14037) doi:10.2196/14037

KEYWORDS

eHealth; cognitive behavioral therapy; health plan implementation; chronic fatigue syndrome; attitudes

Introduction

Chronic fatigue syndrome (CFS) is characterized by severe, persistent, and disabling fatigue. The fatigue is neither explained by the presence of a medical or psychiatric condition nor alleviated by rest. According to the revised Centers for Disease Control and Prevention (CDC) consensus criteria for CFS from 2003, 4 out of the following 8 additional symptoms should be present: problems with concentration and memory, sore throat, tender lymph nodes, headache, muscle pain, multi-joint pain, unrefreshing sleep, and postexertional malaise [1,2].

Cognitive behavioral therapy (CBT) for CFS is aimed at changing behavior and beliefs that maintain symptoms and can effectively reduce fatigue and disability [3,4]. Face-to-face CBT is intensive, requiring 12 to 16 sessions, and therapists need additional training to effectively deliver CBT for CFS [5,6]. Unfortunately, few therapists are trained for CBT for CFS, and therefore, treatment capacity is limited. To overcome this problem, Internet-based CBT (I-CBT) for CFS was developed, which was expected to demand less of therapist resources (ie, therapist time) and also be less burdensome for patients (no need to travel and working at own pace [7]). I-CBT for adult CFS patients was compared with a waiting-list condition in a randomized controlled trial (RCT) in a tertiary CFS treatment center in the Netherlands. It was found to lead to a significant reduction of fatigue and disability while taking approximately 45% less therapist time compared with that of face-to-face CBT (5:23/12:00 hours) [8]. As outcomes appeared less favorable than in face-to-face CBT and not all patients profited, I-CBT was subsequently embedded in stepped care: patients who were still severely fatigued or disabled after I-CBT could step up to face-to-face CBT. Stepped care was compared with care as usual, that was, only face-to-face CBT in a randomized controlled noninferiority trial. Stepped care was effective, and more efficient than care as usual, as it required less therapist time [9].

However, care that has proven to be effective within the context of an RCT, in a tertiary research center, is not necessarily equally effective in routine clinical care [10,11]. In case of CFS, it was found that face-to-face CBT could be provided in mental health care centers (MHCs) with the same magnitude of treatment effect as found in RCTs conducted in tertiary CFS research centers, although not all MHCs reached the benchmark [5,6]. It is not yet known if I-CBT for CFS implemented in routine care is effective. For I-CBT in other disorders, such as depression and anxiety, tinnitus, and irritable bowel syndrome, the first results suggest that it can be successfully implemented in routine clinical care, but more studies are needed [12]. Despite the evidence of the efficacy of electronic health (eHealth), it is incorporated in routine clinical care on a far smaller scale than expected [13,14]. Therefore, it is not yet known how I-CBT can best be embedded in routine clinical care. The first aim of this study is to investigate whether stepped care, comprising I-CBT followed by face-to-face CBT, can be delivered in routine clinical care as effective as in the RCT [9], with respect to treatment outcome.

Second, we are interested in the role of therapist variations on treatment outcomes in routine clinical care. The extent to which differences between the effectiveness of individual therapists explain variance in treatment outcome differs largely over studies. Less influence of therapist variation was associated with therapists being more experienced and the use of treatment manuals [15]. With respect to the treatment of CFS, in a large study in a specialist center, variance in outcome could not be explained by therapist factors. A possible explanation for this finding was that in specialized centers, therapists received the same training and supervision and had the same therapeutic orientation [16]. In a study evaluating the role of the therapist in routine clinical care for CFS (ie, face-to-face CBT), 21% of the variance in treatment outcome was explained by the therapist effect. Attitudes of individual therapists (ie, the attitude toward the use of treatment manuals) were associated with treatment outcome [17].

If therapists' attitudes influence outcome, this provides an opportunity to enhance the efficacy of the treatment, as attitudes can be altered, for example, by experience or training [18,19]. We aimed to investigate if therapists' attitudes toward treatment manuals also influenced outcome in implementation of I-CBT and stepped care for CFS in clinical routine care. Furthermore, we were interested in the role of attitudes of therapists toward eHealth. It was recently found that a positive attitude of the therapists toward eHealth was associated with sharing of more assignments with the therapist by the patient [20]. It was expected that attitudes of therapists influenced treatment outcome in our implementation study.

To answer our research questions, stepped care was implemented in 5 MHCs, with outpatient treatment centers in 9 different cities

spread over the Netherlands. All CFS patients who were referred for CBT were offered I-CBT. If still severely fatigued or disabled after I-CBT, they were offered additional face-to-face CBT. We compared the reduction of fatigue and disability with the benchmark, which was, the effect of stepped care with I-CBT found in an RCT in a tertiary treatment center [9]. Both the effects of the I-CBT (the first step) and the full stepped care model were evaluated. We also explored outcome variations between individual therapists. More specifically, we determined if variation in treatment outcome of stepped care could be explained by therapists' attitudes toward manualized treatment of CFS and the use of eHealth.

Methods

Design

This was an observational study with a pre- and posttreatment study design. Reduction of fatigue severity and level of disabilities were compared with a statistical benchmark, derived from a randomized controlled noninferiority trial for stepped care with I-CBT in a tertiary treatment facility [9] (registered in the Netherlands Trial Register as NTR4809). The latter study included 2 stepped care conditions, of which, the format of therapist feedback during I-CBT differed. In 1 condition, therapists' feedback was given at predefined time points. In the other condition, therapists' feedback was on demand. Both conditions were combined to represent the benchmark in this study (n=242).

MHCs that already offered face-to-face CBT for CFS for at least 1 year were asked to participate in the study. In the participating MHCs, I-CBT was implemented as a first step of stepped care, with additional face-to-face CBT as the second step.

The medical ethical committee of the Radboud university medical center ruled that the study did not fall under the scope of the Medical Research Involving Human Subjects Act (see Multimedia Appendix 1).

Participants

Participating Mental Health Care Centers

Table 1 shows the characteristics of the participating MHCs.

The participating therapists were previously trained to deliver face-to-face CBT, during a 4-day training program in CBT for CFS followed by 2-week supervision for 1 year. All had gained experience in the face-to-face treatment of CFS [5].

In the context of the implementation of I-CBT, therapists had 2 additional training days. The first day focused on delivering I-CBT. The second day was scheduled after the first patients received treatment. The second training day focused on overcoming challenges that were met during delivering I-CBT and the process of stepping up to face-to-face CBT when necessary. Specific for face-to-face CBT after I-CBT is that although it follows a treatment manual [21], the starting point differs for each patient. To tailor the CBT to the patients' process, therapists were trained to (1) identify what was already achieved during I-CBT and what was needed to further improve and (2) motivate patients to actually step up to face-to-face CBT. Therapists were trained to recognize and modify reduced expectations of the patient toward face-to-face CBT, after limited results during I-CBT.

Table 1. Characteristics of the participating mental health care centers.

Mental health care center	Treatment sites, n	Location site(s) in the Netherlands	Therapists
PsyQ Parnassia Groep	4	Central and west (the 4 largest cities of the Netherlands)	2-3 psychologists per site
PsyQ Lentis	2	Northeast	2 psychologists per site
PsyQ MET ggz	1	South	2 psychologists
GGNet	1	East	3 psychologists
GGz Westelijk Noord-Brabant	1	Central southwest	I-CBT[a]: 4 psychiatric nurses; face-to-face CBT[b]: 3 psychologists

[a]I-CBT: internet-based cognitive behavioral therapy.

[b]CBT: cognitive behavioral therapy.

Participating Patients

All adult patients referred for the treatment of CFS could participate if the following criteria were met: (1) a physician had concluded that the patient suffered from severe and disabling fatigue not explained by a known somatic or psychiatric condition; (2) the 2003 CDC consensus criteria for CFS were met (ie, severe, disabling fatigue was present, lasting for at least 6 months, accompanied by at least 4 out of 8 additional symptoms) or patients met criteria for *idiopathic chronic fatigue* (ICF) *syndrome* (ie, reported severe and persistent fatigue but did not meet all CDC criteria, <4 additional symptoms, or less impact on daily functioning [22]). Inclusion criteria for both patient groups were the presence of severe and persistent fatigue, as indicated by a score of ≥35 on the Checklist Individual Strength (CIS) Fatigue Severity Subscale [23], limitations in functioning according to the Short Form-36 (SF-36) Physical or Social Scale <65 [24], or both >65 but the patient was limited in daily functioning according to a clinical interview, for example, worked less; and (3) the patient had computer and internet access. There were no specific exclusion criteria.

Intervention

All patients were offered I-CBT as a first step of treatment. If patients met step-up criteria after 6 months of I-CBT (still severely fatigued as indicated by CIS fatigue severity ≥35 or limited in functioning as indicated by SF-36 Physical or Social Functioning Scale ≤65), face-to-face CBT was offered.

Both forms of CBT were based on a treatment manual [21] that has been used in RCTs testing the efficacy of I-CBT and face-to-face CBT for CFS [25,26]. CBT comprises interventions aimed at changing behavior and cognitions that maintain CFS symptoms. It starts with setting concrete goals in terms of activity, which when reached, imply recovery (ie, no longer severely fatigued and disabled). Patients learn to establish a fixed sleep-wake cycle to recognize and modify dysfunctional cognitions and redirect their focus on symptoms to other matters. After this, patients start with a graded (physical) activity program, usually walking or cycling. The activity program is tailored to the patient, based on their activity pattern. The increase in activity is time contingent, irrespective of symptoms. In the same manner, social and mental (eg, reading) activities are increased and personal goals are attained.

The efficacy of I-CBT was tested in an RCT [8], and it is described in detail elsewhere [27]. The intervention comprises 7 modules corresponding to the different elements of the face-to-face protocol. After the first module is finished (*getting started and goal setting*), the following 5 modules become accessible. These modules were as follows: *regulate sleep-wake cycle*, *helpful beliefs about fatigue*, *how to communicate with others about CFS*, and *gradually increasing my activity*. When the sixth module *reaching my goals step by step* is finished, the seventh module opens (*evaluation and the future*). The duration of the modules differed, and the patient could work through them at his or her own pace for 6 months. Therapists were instructed to provide feedback weekly in the first month and at least fortnightly in the following 5 months. Patients were sent reminders if they did not report on their progress according to the aforementioned schedule. Therapists could respond to an assignment the patient completed or send an email via the platform. Feedback was aimed at helping the patients change their behavior and cognitions according to the principles they learned in the modules [27].

Accessibility differed between MHCs. A total of 3 MHCs used the same platform as the tertiary treatment center, and 2 MHCs incorporated the content of the intervention into their own eHealth portal. All used the same treatment content.

After 6 months of I-CBT, patients were invited for an face-to-face evaluation session and offered additional face-to-face CBT when they still met the aforementioned step-up criteria. The face-to-face CBT was delivered according to the treatment manual, although the starting point was tailored to the needs of the patient, as some cognitions and behaviors already changed during I-CBT. During the evaluation session, it was examined which cognitions and behavior remained dysfunctional and would be the focus of the additional face-to-face CBT. The number of CBT sessions could vary, and the expected maximum duration was 6 months. The face-to-face CBT therapist was preferably the same therapist who delivered I-CBT. In 1 MHC, this was not possible, as I-CBT was provided by a psychiatric nurse, whereas the face-to-face CBT was provided by a psychologist. Information on the course of the treatment was given to the psychologist.

Measures

Before the start of the study, all therapists were asked to complete questionnaires assessing their attitudes toward the manualized treatment and the use of eHealth.

Patient data were collected by therapists. Some MHCs used digital questionnaires, some used pencil and paper versions of the outcome measures. The therapists gave each participant a unique research number and entered the relevant data into a spreadsheet, listed by research number and without personal information, except for age and sex. The spreadsheet was accessible to the researcher.

For patients who received the full stepped care, outcomes were measured at baseline, after I-CBT (6 months after start), and directly after face-to-face CBT (posttreatment assessment). For patients who received I-CBT only, the outcome directly after I-CBT was used as posttreatment assessment.

Fatigue Severity

Fatigue severity was measured with a fatigue questionnaire, the 20-item CIS [28], that measured different aspects of fatigue. The Fatigue Severity subscale is used to assess the level of fatigue and comprises 8 items, scored on a 7-point Likert scale (range 8-56, higher scores indicate more severe fatigue). The CIS has proven to be reliable (Cronbach alpha for fatigue severity subscale ranges between .69 and .94 [23,28]) and valid and has been used extensively in CFS research as outcome measure [28].

Functioning

Physical and social functioning was measured with the Medical Outcomes Survey SF-36 [24], a reliable and valid instrument to measure health status [29]. Cronbach alpha of the Dutch version is .92 for the Physical Functioning Subscale and .71 for the Social Functioning Subscale [30]. The Physical Functioning Subscale assesses physical functioning with 10 items. Scores on the scale range from 0 to 100, with higher scores indicating better physical functioning. The Social Functioning Subscale assesses impairment in social functioning with 2 questions scored. Total scores range from 0 to 100; higher scores indicate better social functioning.

Impairment in daily functioning was measured with the Work and Social Adjustment Scale (WSAS [31]). This scale assesses functioning at work, in home management, and in social and leisure activities, using 5 items, on a scale ranging from 0 to 8 (range of total score is 0-40, with higher scores indicating more impairment). The Dutch version of the WSAS is validated in CFS patients; Cronbach alpha is .89 [32].

Additional Symptoms

The number of additional CDC symptoms [1] was registered during the interview. Some therapists used a pen-and-paper checklist that systematically checked existing symptoms and a minimal duration of 6 months. The symptom maximum was 9,

as concentration and memory problems were recorded separately.

Therapists' Attitudes

Therapists' attitudes toward the use of treatment manuals were assessed with a questionnaire developed by Addis et al [33] to measure attitudes of psychologists. It measures 2 constructs: (1) *Positive Outcome* (7 items), which reflects the attitude that manuals can contribute positively to treatment outcome (Cronbach alpha=.93) and (2) *Negative Process* (10 items), which reflects the attitude that the use of treatment manuals negatively influence the treatment process (Cronbach alpha=.80). A Dutch version was used, with a 6-point scale scored from 0 to 5 (positive outcome range 0-35, negative process range 0-50) [17].

Attitudes toward the use of eHealth were measured with an 18-item version of the eHealth attitude list [19]. The scale *Possibilities of eHealth* contains 7 items (5-point Likert scale, range 7-35), and higher scores reflect the attitude that eHealth can be valuable. The scale *eHealth Negative Effects* contains 9 items (5-point Likert scale, range 9-45), and higher scores represent the attitude that eHealth poses a threat to the therapy process. The scale *Computer Competence* contains 2 items (5-point Likert scale, range 2-10), and a higher score indicates that the therapist feels competent in using computers. Structural and internal validity and internal consistency are good (Cronbach alpha between .83 and .89) [19,20].

Statistical Analyses

Analyses were conducted after imputation of missing primary outcomes after I-CBT and missing primary and secondary outcomes at posttreatment (after face-to-face CBT or after I-CBT for patients who did not receive face-to-face CBT), using multiple imputations and assuming data were missing at random. A total of 20 imputations were performed. CIS, SF-36 Physical functioning and Social functioning, and total score on the WSAS at baseline and posttreatment were entered as predictors and variables to impute.

In the dataset used for the statistical benchmark [9], the CIS and the SF-36 Physical functioning were imputed in the same manner. For the purpose of this study, WSAS and SF-36 Social functioning scale were imputed likewise in the dataset.

Imputation and statistical analyses were performed with IBM SPSS version 22.

Treatment Effects

Treatment effects were tested with paired samples *t* tests for each outcome measure. To answer the primary research question, it was determined whether the change score between baseline and posttreatment assessment fell within the 95% CI of the change scores found in the RCT performed in the tertiary CFS research center (n=242) [9].

To specifically explore the efficacy of I-CBT implemented in the MHCs, fatigue severity before and after I-CBT was compared using a paired samples *t* test, and the change score in fatigue severity of I-CBT was compared with the benchmark of I-CBT from the RCT.

Uncontrolled effect sizes (within group Cohen *d*) were calculated for the CIS fatigue severity subscale, SF-36 physical and social functioning, and for the WSAS total score [34]. This was done by dividing the difference between the mean at baseline and postassessment by a pooled standard deviation ($\sqrt{[SD_{pre}^2+SD_{post}^2]}/2$). CIs were calculated following Hunter and Schmidt [35].

A sensitivity analysis was performed, in which missing data on CIS fatigue severity were not imputed but replaced by the maximum score (56). This was done with the assumption that patients who had no postassessment deteriorated. In the benchmark study, there were no missing CIS values.

Proportion of Patients With Clinically Significant Improvement in Fatigue Severity

A clinically significant improvement in fatigue severity was defined as a statistically Reliable Change Index (RCI) of >1.96 SD in CIS fatigue severity [36], in combination with a CIS fatigue severity score of <35 on postassessment. The score of <35 indicates that the patient is no longer severely fatigued [28]. The reliability of the CIS used in the RCI calculation was 0.88 [9,23]. An RCI of >1.96 SD means it can be assumed with a confidence of 95% that the improvement in CIS fatigue severity is not caused by unreliability of the measure but represents a true change.

Subgroup Analyses

To facilitate comparison with the benchmark study, in which patients with ICF syndrome were not included, change scores were calculated for the subgroup of patients in implemented stepped care who did meet the CDC criteria for CFS.

Therapists' Attitudes and Treatment Outcome

A mixed-models approach was planned to investigate to what extent variance in treatment effect could be explained by the therapists and their attitudes toward I-CBT and the use of treatment manuals. If the number of patients per therapist would be too low, variations in treatment outcome between therapists would be explored by dividing the therapists in 2 groups using a median split based on CIS fatigue change scores. Reduction in fatigue severity in patients of both groups would be compared using a *t* test.

To explore if therapists' attitudes toward I-CBT and treatment manuals were related to treatment outcome, correlations were calculated between the therapists' attitude subscale scores and the mean change score in fatigue severity of their patients. When no more than 1 item on an attitude subscale was missing, the missing value was replaced with the mean score on that subscale.

Results

Overview

Data of all 100 participants were analyzed. Postassessment data of the primary outcome measure (CIS fatigue severity) of 87 participants (87/100, 87.0%) was present.

From October 2014 to December 2016, 125 patients were referred for treatment for CFS (Figure 1). Of them, 100 were

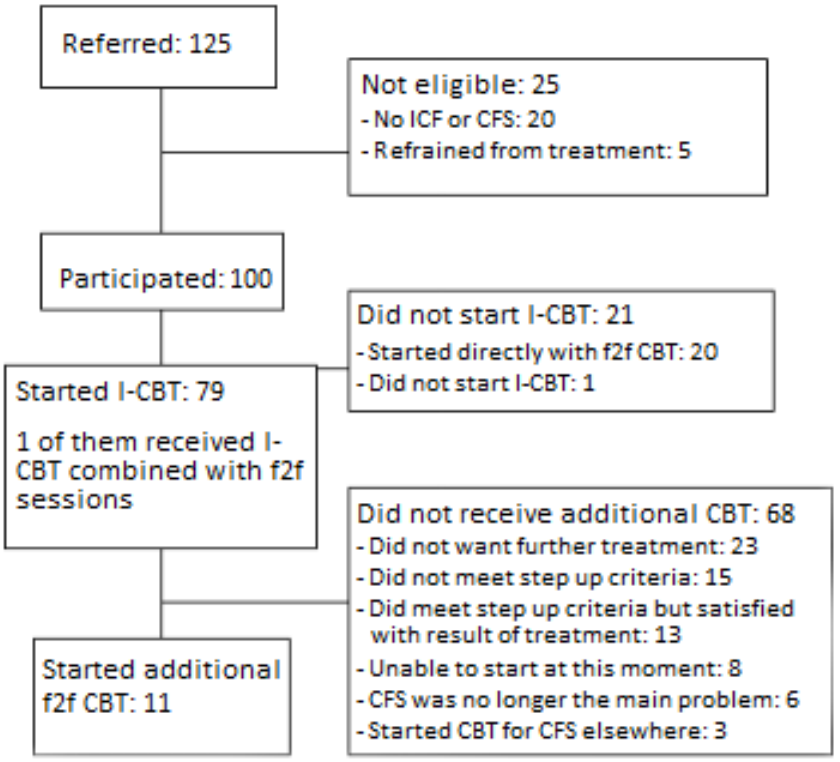
eligible to enter the study and were included, 20 had no CFS (were not severely fatigued or had another diagnosis that explained the presence of fatigue), and 5 did not want treatment.

Out of the 100 eligible patients, 73 met all CDC criteria for CFS, 10 patients had <4 additional symptoms, for 14 patients the number of additional symptoms was unknown, and 3 patients were not severely impaired in functioning according to the SF-36 (both social and physical functioning ≥65) but reported severe impairment during the clinical interview.

In total, 79 patients started with I-CBT as intended, whereas 20 patients started directly with face-to-face CBT (20%). For 7 patients, it was reported that the patient preferred face-to-face CBT; for the others, reasons were not reported. In the benchmark study, 14 patients (14/242, 5.8%) who intended to start with I-CBT, started directly with face-to-face CBT [9], whereas 119 patients (119/766, 15.5%) eligible to enter the trial refused because they preferred face-to-face CBT. In the implementation study, 1 patient did not start treatment. For 24 patients (24/79, 30%), the therapist assumed dropout during I-CBT (no response and no new log-ins observed, or the patient explicitly reported to have stopped).

After I-CBT, 15 patients (15/79, 19%) did not meet the step-up criteria (were no longer severely fatigued or disabled). For 16 patients (16/79, 20%), it is not known whether step-up criteria were met, as post–I-CBT assessment scores were missing. The other 48 patients (48/79, 61%) with posttreatment data met step-up criteria. In total, 11 patients stepped up to face-to-face CBT (11/48, 23%). Reasons for not stepping up are given in the flowchart (Figure 1). In the benchmark study, 172 patients (172/242, 71.1%) met step-up criteria after I-CBT. Of them, 85 (49.4%) stepped up to face-to-face CBT [9].

Figure 1. Flowchart. CBT: cognitive behavioral therapy; CFS: chronic fatigue syndrome; f2f: face-to-face; I-CBT: internet-based cognitive behavioral therapy; ICF: idiopathic chronic fatigue.



Patient Characteristics

Baseline characteristics are shown in Table 2. The proportion of female patients was larger than that in the stepped care arms of the RCT [9], with which the data were compared, which had a relatively low proportion of females [8]. Age in years, fatigue severity, and physical functioning did not significantly differ between this study sample and the benchmark sample (see Table 2). Social functioning and impairment in daily functioning were significantly worse in the MHC sample, whereas patients in the benchmark sample reported significantly more additional CDC symptoms.

Table 2. Baseline characteristics of this study and the benchmark study.

Baseline characteristic	MHC[a] sample—stepped care in routine clinical care	Benchmark sample—stepped care in RCT[b] (n=242)	Difference between samples			
			Chi-square (*df*)	*t* test (*df*)	Mann-Whitney *U* test	*P* value
Proportion female, n (%)	78 (78)	147 (60.7)	9.4 (1)[c]	—[d]	—	.002
Age (years; n=100), mean (SD)	37.4 (11.9)	36.9 (12.5)	—	−0.345 (340)	—	.73
Fatigue severity (CIS[e]; n=100), mean (SD)	49.6 (5.2)	50.5 (4.9)	—	1.580 (340)	—	.12
Physical functioning (SF-36[f]; n=98), mean (SD)	60.3 (21.3)	61.4 (19.7)	—	0.455 (338)	—	.65
Social functioning (SF-36; n=96), mean (SD)	37.2 (23.8)	44.0 (23.7)	—	2.365 (336)	—	.02
Impairment daily functioning (WSAS[g]; n=92), mean (SD)	25.6 (6.5)	23.2 (6.7)	—	−2.957 (332)	—	.003
Number of additional CDC[h] symptoms (0-9; n=86), median (IQR[i])	6 (3)	7 (2)	—	—	7703	<.001

[a]MHC: mental health care center.

[b]RCT: randomized controlled trial.

[c]n=342.

[d]Not applicable.

[e]CIS: Checklist Individual Strength.

[f]SF-36: Short Form-36.

[g]WSAS: Work and Social Adjustment Scale.

[h]CDC: Centers for Disease Control and Prevention.

[i]IQR: interquartile range.

Treatment Effect

As shown in Table 3, patients significantly improved on all outcomes. Compared with the benchmark, the decrease in fatigue severity (CIS) is lower in the MHC sample, as the change score falls outside the 95% CI of the change in the benchmark study. For physical functioning and social functioning, the change scores fall within the CI of the benchmark. For limitations measured with the WSAS, the change score is above the CI of the benchmark.

For the sensitivity analysis, the CIS fatigue severity postassessment scores of all 13 patients with missing data were replaced with the maximum score of 56. This reduced the fatigue change score to 10.1, which was still a significant improvement (t=7.2, P<.001), that falls outside this CI.

Proportion of Patients With Clinically Significant Improvement in Fatigue Severity

Data were missing for 13 patients, and it was assumed that they did not show a significant change in fatigue severity. Of the 100 patients, 37 (37/100, 37%) had a reliable and clinically significant improvement in fatigue severity and were no longer severely fatigued. In the stepped care conditions in the tertiary center, no CIS data were missing and 110 (110/242, 45.5%) had a clinically significant improvement in fatigue. The difference between the 2 improvement rates was not significant; χ^2_1(N=342)=2.1 and P=.15.

Subgroup Analyses

When restricting the analyses to patients who met CDC criteria for CFS (n=73), scores on all outcomes were significantly improved (Table 3). The change scores on all outcomes were larger than that in the total group. All scores fall within the CIs of the tertiary treatment center, except for the change score in limitations (WSAS), which falls outside the CI of the benchmark. A post hoc analysis of variance comparing the CFS group with the ICF group (n=14) and the group for which, because of missing data, it was unknown whether the diagnosis was CFS or ICF (n=13), showed a group effect. CFS patients had a significantly larger reduction in fatigue severity than the patients with diagnosis unknown. Differences between the ICF group and both other groups were not significant.

The proportion of patients with a reliable and clinically significant improvement in fatigue severity was 29 out of 73 (40%). Data of 11 patients were missing, and no improvement was assumed for them.

Table 3. Treatment effect.

Outcome measure and selected group	n (%)	Treatment		Change score (95% CI)	*t* test	*P* value	Effect size (*d*)
		Pre	Post				
Fatigue severity (CIS[a])							
MHC—total group[b]	100 (13)	49.6	37.0	12.6 (9.7 to 15.5)	8.5	<.001	1.14 (0.84-1.44)
MHC—CFS only[c]	73 (15)	50.2	35.9	14.4 (11.0 to 17.8)	8.3	<.001	1.31 (0.95-1.67)
Benchmark[d]	242 (0)	50.5	35.6	14.9 (13.2 to 16.5)	17.8	<.001	1.47 (1.27-1.67)
Physical functioning (SF-36[e])							
MHC—total group	100 (32)	60.1	73.5	−13.4 (−18.1 to −8.7)	−5.6	<.001	0.62 (0.34-0.91)
MHC—CFS only	73 (33)	59.0	75.4	−16.4 (−21.9 to −10.9)	−5.9	<.001	0.76 (0.42-1.10)
Benchmark	242 (5)	61.4	76.6	−15.2 (−17.6 to −12.8)	−12.3	<.001	0.71 (0.53-0.90)
Social functioning (SF-36)							
MHC—total group	100 (34)	37.0	57.5	−20.4 (−27.9 to −12.9)	−5.4	<.001	0.73 (0.44-1.01)
MHC—CFS only	73 (34)	33.7	59.3	−25.6 (−34.0 to −17.2)	−6.0	<.001	0.96 (0.61-1.30)
Benchmark	242 (5)	44.0	65.5	−21.5 (−25.2 to −17.8)	−11.4	<.001	0.84 (0.65-1.02)
Limitations (WSAS[f])							
MHC—total group	100 (40)	26.1	15.8	10.3 (7.8 to 12.7)	8.3	<.001	1.08 (0.78-1.38)
MHC—CFS only	73 (40)	26.5	15.0	11.4 (8.6 to 14.3)	7.9	<.001	1.24 (0.88-1.59)
Benchmark	242 (14)	23.2	14.6	8.6 (7.4 to 9.8)	14.0	<.001	0.99 (0.81-1.18)

[a]CIS: Checklist Individual Strength.

[b]MHC (mental health care center)—total group: all 100 participants, regardless of meeting Centers for Disease Control and Prevention criteria for chronic fatigue syndrome.

[c]CFS (chronic fatigue syndrome) only: subgroup of 73 participants that met Centers for Disease Control and Prevention criteria for chronic fatigue syndrome.

[d]Benchmark: patients who were allocated to the stepped care arms of the randomized control trial, all meeting Centers for Disease Control and Prevention criteria for chronic fatigue syndrome.

[e]SF-36: Short Form-36.

[f]WSAS: Work and Social Adjustment Scale.

Of the 80 patients who intended to start I-CBT, 64 had completed the post–I-CBT assessment, and 16 CIS fatigue severity scores were imputed. The mean CIS fatigue score after I-CBT was 36.7, which was on average 13.2 points lower than that at the preassessment (95% CI 9.8-16.5; t=7.8; P<.001). This change score falls within the 95% CI of the benchmark from the patients who followed I-CBT in tertiary treatment center (95% CI 11.1-14.2). A sensitivity analysis was performed by replacing all missing CIS fatigue scores post–I-CBT with the maximum fatigue score (56). This resulted in a change score of 8.7 (95% CI 5.6-11.7; t=5.6; P<.001), which fell below the benchmark.

Therapists' Attitude and Treatment Outcome of Fatigue Severity

In total, 25 therapists participated in the study. Of them, 15 therapists treated at least one patient and completed attitude questionnaires, 5 treated at least one patient but did not complete these questionnaires, and 5 completed the questionnaires but did not treat a patient. The number of patients per therapist varied from 1 to 18.

The 20 therapists who treated patients were ranked based on the median fatigue severity change score of their patients. The 10 therapists with the lowest median change score treated 43 patients, with a mean change score in fatigue severity of 9.0 points. The 10 therapists with the highest median change scores treated 57 patients who had a mean change score of 15.7. The difference between these means was significant (see Table 4). No differences between the 2 groups were found in therapists' attitudes, except for the *Computer Competence* scale. The therapists with higher median fatigue change scores had significantly higher scores on *Computer Competence*. Furthermore, differences on baseline characteristics (fatigue, physical functioning, social functioning, number of additional CDC symptoms, level of limitations, age, and sex) of the patients treated by both therapist groups were compared using t tests and chi-square tests, and no difference was significant (P=.07-.93).

Table 4. Therapists attitude and treatment outcome (data of all 20 therapists who completed the questionnaires provided and of whom, 5 had not treated a chronic fatigue syndrome patient during the study).

Variable	All therapists	Therapists		Statistical difference between therapists with high and low median	
		With high median	With low median	*t*	*P* value
Patients treated[a], n (%)	100 (87)	57 (48)	43 (39)	—[b]	—
Change score in fatigue severity, mean	12.7	15.7	9.0	2.42	.02
Attitude eHealth[c], mean (SD)					
Possibilities of eHealth scale (range 9-45)	28.0 (2.6)	28.6 (2.5)	25.8 (2.5)	−2.05	.06
eHealth Negative Effect Scale (range 7-35)	25.2 (5.9)	25.6 (6.3)	26.8 (4.6)	0.38	.71
Computer Competence Scale (range 2-10)	8.1 (1.7)	8.7 (1.3)	7.0 (1.4)	−2.28	.04
Attitude manualized treatment, mean (SD)					
Positive outcome (range 0-35)	24.9 (4.1)	24.8 (4.7)	24.8 (4.1)	0.00	>.99
Negative process (range 0-50)	14.6 (7.4)	14.9 (8.1)	13.8 (7.4)	−0.25	.80

[a]Number of patients with complete data.

[b]Not applicable.

[c]eHealth: electronic health.

The correlation between the fatigue severity change score per therapist and the attitude subscales were as follows: .363 for *Possibilities of eHealth* (*P*=.18), −.092 for *eHealth Negative Effect* (*P*=.75), .271 for *Computer Competence* (*P*=.33), .119 for *Positive Outcome* of treatment manuals (*P*=.67), and −.186 with *Negative Process* of treatment manuals. None of the correlations were significant (*P*=.51).

Discussion

Principal Findings

This is one of few studies that implemented and evaluated eHealth in routine clinical care. It shows that I-CBT embedded in stepped care for CFS can lead to a significant reduction in fatigue severity and limitations in routine clinical care. The outcomes were compared with those of the same treatment format delivered in an RCT, in a tertiary treatment center for CFS, the benchmark. Outcomes for limitations in functioning were similar in both settings or even better in the implemented care. Nevertheless, the decrease in fatigue severity was smaller in the MHCs. Compared with the benchmark, relatively fewer patients had a clinically significant improvement in fatigue severity (reliable change in fatigue and no longer severely fatigued) after stepped care, but the difference was not statistically significant. It should be noted that the improvement rate in the benchmark is also lower than what was previously found in the tertiary treatment center in an face-to-face treatment [3]. Although these studies are not entirely comparable, this suggests that there is probably room for improvement in the delivery of this treatment by MHCs in routine clinical care. We will elaborate on this in the future directions section.

We found that I-CBT, the first step of stepped care, can be delivered effectively in routine clinical care. It led to a significant reduction of fatigue severity. Patients who received I-CBT in the MHCs did not profit less from the intervention than those in the tertiary treatment center. However, in the RCT, there were no missing values, whereas in the MHCs, post–I-CBT data of one-fifth of the patients had to be imputed. As treatment outcome for research dropouts may not be the most favorable, it is possible that we overestimated the efficacy of implemented I-CBT. Our sensitivity analysis, in which we assumed that these patients are maximally fatigued, showed a smaller reduction of fatigue in the MHCs, falling below the benchmark.

An important finding is that one-fifth of the patients did not start with I-CBT but directly commenced with face-to-face CBT. In the benchmark study, almost all patients started with I-CBT. Nevertheless, we can assume that patients who did not want I-CBT did not participate in an RCT testing it, which is a general limitation of testing I-CBT in an RCT. It is known that 15.5% of the patients eligible to enter the RCT refused because they preferred face-to-face CBT. An important issue for the interpretation of our results is that we cannot know how many patients who started directly with face-to-face CBT would have improved in I-CBT. They were probably not a random sample of all eligible patients. If this subgroup, for example, already expected to profit less from I-CBT, not including them might have had led to inflated results of I-CBT. Likewise, if these were mainly patients of therapists with little confidence in I-CBT, who, therefore, less convincingly offered it, they might have been better off in face-to-face CBT, which also would have led to inflated results of I-CBT after implementation.

According to therapist reports, about 30% of patients who started I-CBT dropped out during I-CBT. Although this figure is informative, it should be noted that therapist reports may not be the most reliable indication of dropout in I-CBT [9]. Furthermore, of the patients who were still severely fatigued and impaired after implemented I-CBT, about 1 in 5 stepped up to face-to-face CBT. In comparison, in the tertiary treatment center, approximately half of the patients stepped up when still fatigued or impaired after I-CBT. In the RCT, it was found that stepping up generally led to additional therapy gains. This could partly explain why in the RCT, patients had a larger decrease

in fatigue severity: relatively more patients in the RCT received both I-CBT and face-to-face CBT. The problem of not starting and not stepping up is common in stepped care [37,38].

Another study [39] compared the outcome of stepped care for CFS in routine clinical care (ie, MHC) with the outcome of CBT in a tertiary treatment center. The first step was CBT using a self-help booklet with therapist guidance via email, and the second step was face-to-face CBT. Compared with outcomes in the tertiary treatment center, the MHC showed a lower reduction of fatigue severity and, contrary to our study, also a lower increase in physical functioning. Interestingly, previous studies found that the guided self-help, as well as the face-to-face CBT, could be delivered as effectively in the MHC as in the context of an RCT [5,6,26]. However, in combination with face-to-face CBT, the results were less positive. This suggests that there is something specific to stepped care that makes it more difficult to deliver in routine clinical care than face-to-face CBT, I-CBT, or guided self-help alone [39]. Several explanations could be considered. Delivering face-to-face CBT after a minimal intervention with limited results may demand more treatment experience from therapists than starting face-to-face CBT from the start, also because the patients in need of face-to-face CBT may be relatively complex to treat [6]. In our study, therapists were more experienced, and training had paid special attention to these difficulties specific to stepped care. Still, it seems that despite the training, many patients did not step up after unsuccessful I-CBT.

In the comparison of routine clinical care and care in a tertiary treatment center, it is important to consider possible differences between patients of both settings that may influence outcome. For CFS, it was found in routine clinical care that samples more often comprised patients with psychiatric comorbidity [26], which might negatively influence outcome [40]. Furthermore, we anticipated the inclusion of patients with ICF, because MHCs treat these as well, as they too benefit from CBT [22]. Interestingly, our study showed when only selecting patients who met CDC criteria for CFS, the reduction of fatigue in the MHCs did fall within the benchmark (whereas when selecting the total group, this fell below the benchmark). Our post hoc analysis showed that patients with ICF did not profit less than patients with CFS, but the group of patients for whom the diagnosis (ICF or CFS) was unknown had significantly lower treatment outcome.

We evaluated the influence of therapist factors on treatment outcome. In accordance with a former study investigating the role of the therapist on treatment outcome after implementation [17], we found considerable variance in outcome between therapists. We expected that therapists' attitudes toward eHealth and manualized treatment would influence treatment outcome. Unfortunately, because of the limited number of patients per therapist, we were unable to perform the multilevel analysis we planned, which is a limitation of the study. The correlations between therapists' attitudes and fatigue reduction were in the expected direction, but none was significant. Replication with more patients per therapist is needed, as it is likely that we lacked statistical power because of the small sample size.

Strengths and Limitations

A limitation of the study is that the moments on which feedback was given during the I-CBT were not exactly the same as in the benchmark study. In the latter study, patients had received either I-CBT with feedback on prescheduled moments (*protocol-driven therapist feedback*) or I-CBT with feedback only when the patient asked for it (*on demand*) [8,9]. As both treatment arms did not differ in outcome, both were combined to calculate the benchmark. In this study, therapists were advised to give feedback according to the feedback schedule used in the *protocol-driven* condition of the benchmark study, but during supervision, it became clear that therapists did not always follow the schedule.

Furthermore, step-up criteria differed slightly. Additional face-to-face CBT was offered in both studies when patients were still severely fatigued and impaired following I-CBT. In the benchmark study, the Sickness Impact Profile 8 was used to assess the level of disability, but it was too lengthy for use in the MHCs. Therefore, the SF-36 was used to assess whether severe limitations in physical or social functioning were present. Although both instruments are used to assess limitations in functioning, it is possible that a different subgroup of patients were selected to step up using the 2 measures.

Another limitation is that we do not have information on usage of the program, either for patients or for therapists. For example, it is not recorded when therapist feedback was provided. Furthermore, the duration of the face-to-face CBT and the number of sessions was not registered.

Finally, the data collection was, to warrant patient privacy, done by the therapists, and questionnaires were sometimes scored by hand. This procedure may have reduced the reliability of data collection.

There were some important strengths as well. First, our study not only evaluated I-CBT in routine clinical care but also shed light on how I-CBT can be embedded in routine care. Second, by comparing the results with a benchmark, we were able to put the results in perspective. Finally, we included multiple treatment centers, across the Netherlands, which contributed to ecological validity of our study.

Future Directions

The success of I-CBT and stepped care in routine clinical care can probably be improved by aiming at avoiding dropout and increasing the numbers of patients stepping up. To avoid dropout, it would be important to know why patients dropped out. Information on usage of the program would be helpful. If, for example dropout occurs more during 1 treatment module than another, that specific module could be improved. It is also important to know how aspects of therapist guidance (quantity and perceived quality) influenced dropout to improve the therapist training.

Likewise, to encourage stepping up, it is important to know what prevented patients from stepping up. The main reasons given were that patients did not want CBT anymore or were unable to start at that moment. More knowledge about reasons for not wanting CBT is needed to develop strategies to increase

the number of patients who step up. If, for example, the long duration of treatment is a reason, one could evaluate the effect of I-CBT sooner, for example, after 3 or 4 months, to step up earlier. This would shorten the duration of the total treatment and may help avoid the loss of motivation to step up [9].

Furthermore, other options to embed I-CBT in routine care should be considered to increase the number of patients who profit from treatment—by offering matched care, for example. Unfortunately, it is not known which patients profit from I-CBT and which patients would profit more from face-to-face CBT. Future research should search for predictors of the outcome of different treatment formats or determine if following patient preference leads to better results. Therapists and patients may be well capable of resolving together what treatment form would work best for the patient. I-CBT and face-to-face CBT are both based on the same treatment principles and protocol, and it is mainly the form of communication that differs. It would be interesting to compare CBT offered as stepped care with preferred care, that is, either I-CBT or face-to-face CBT, depending on the preference of the patient. It was found in a meta-analysis that treatment outcome is higher when the patient receives the treatment of preference [41].

Finally, it is important to also investigate long-term outcome of implemented stepped care. This may be difficult to achieve in an observational multicenter study but would be a valuable contribution. A recent study showed that positive outcome for CFS after CBT is only partially maintained at long-term follow-up up to 10 years after treatment, as 37% still had fatigue within normal ranges [3]. Although efforts should be made to increase this proportion, it confirms the notion that one can recover from CFS and maintain the gains. We cannot assume that the long-term outcome for face-to-face CBT is the same as I-CBT/stepped care. It may, for example, be possible that the interventions at the end of the therapy (goal reaching, evaluation, and preparing the future) are important to retain the accomplishments, although probably a larger proportion of patients did not reach these modules.

Conclusions

This is, to our knowledge, the first study to evaluate I-CBT for CFS, embedded in stepped-care, in routine clinical care. I-CBT was as effective as in tertiary care and could be embedded in routine clinical care where additional face-to-face CBT was offered if needed. Increasing the number of patients who step up after I-CBT is the most important remaining issue for implementation of stepped care.

Acknowledgments

This study was a part of the project *Grip op klachten* (Master Your Symptoms). Master Your Symptoms aims at developing internet-based tools to improve health care for patients with medically unexplained somatic symptoms. This project was funded by Innovatiefonds zorgverzekeraars (Innovation Fund Health Insurances). For this specific study, it was used to pay the salary of MWS. The authors would like to thank the following MHCs for participation: PsyQ Parnassia Groep, PsyQ Lentis, PsyQ MET ggz, GGNet, and GGz Westelijk Noord-Brabant. They would also like to thank all patients for their participation; all therapists involved for delivering CBT and data collection; C van Ginkel, L Lubbers, M Deij, D Schurer, E Alberts, A Janse, and E Klerks for their help with data collection; R Donders and P Nieuwkerk for statistical advice; and J Rosmalen for feedback during the design of the study and her comments on a draft of the paper.

Authors' Contributions

The study was designed by HK, MW, MWS, and AvD. MWS contributed to data acquisition and analysis. Data interpretation was performed by HK, MW, MWS, and AvD. All authors were involved in interpretation of results and in reviewing and revising the paper.

Conflicts of Interest

This project was funded by Innovatiefonds zorgverzekeraars (Innovation Fund Health Insurances). For this specific study, it was used to pay the salary of MWS. HK receives royalties for the CBT for CFS treatment manual that was used in the study.

Multimedia Appendix 1
Decision letter from the medical ethical committee.
[PDF File (Adobe PDF File)101 KB - jmir_v21i10e14037_app1.pdf]

References

1. Reeves WC, Lloyd A, Vernon SD, Klimas N, Jason LA, Bleijenberg G, International Chronic Fatigue Syndrome Study Group. Identification of ambiguities in the 1994 chronic fatigue syndrome research case definition and recommendations for resolution. BMC Health Serv Res 2003 Dec 31;3(1):25 [FREE Full text] [doi: 10.1186/1472-6963-3-25] [Medline: 14702202]
2. Fukuda K, Straus SE, Hickie I, Sharpe MC, Dobbins JG, Komaroff A. The chronic fatigue syndrome: a comprehensive approach to its definition and study. International chronic fatigue syndrome study group. Ann Intern Med 1994 Dec 15;121(12):953-959. [doi: 10.7326/0003-4819-121-12-199412150-00009] [Medline: 7978722]

3. Janse A, Nikolaus S, Wiborg JF, Heins M, van der Meer JW, Bleijenberg G, et al. Long-term follow-up after cognitive behaviour therapy for chronic fatigue syndrome. J Psychosom Res 2017 Jun;97:45-51. [doi: 10.1016/j.jpsychores.2017.03.016] [Medline: 28606498]
4. Castell BD, Kazantzis N, Moss‐Morris RE. Cognitive behavioral therapy and graded exercise for chronic fatigue syndrome: a meta‐analysis. Clin Psychol: Sci Pract 2011;18(4):311-324. [doi: 10.1111/j.1468-2850.2011.01262.x]
5. Scheeres K, Wensing M, Knoop H, Bleijenberg G. Implementing cognitive behavioral therapy for chronic fatigue syndrome in a mental health center: a benchmarking evaluation. J Consult Clin Psychol 2008 Feb;76(1):163-171. [doi: 10.1037/0022-006X.76.1.163] [Medline: 18229994]
6. Wiborg JF, Wensing M, Tummers M, Knoop H, Bleijenberg G. Implementing evidence-based practice for patients with chronic fatigue syndrome. Clin Psychol Psychother 2014;21(2):108-114. [doi: 10.1002/cpp.1827] [Medline: 23229956]
7. Gratzer D, Khalid-Khan F. Internet-delivered cognitive behavioural therapy in the treatment of psychiatric illness. Can Med Assoc J 2016 Mar 1;188(4):263-272 [FREE Full text] [doi: 10.1503/cmaj.150007] [Medline: 26527829]
8. Janse A, Worm-Smeitink M, Bleijenberg G, Donders R, Knoop H. Efficacy of web-based cognitive-behavioural therapy for chronic fatigue syndrome: randomised controlled trial. Br J Psychiatry 2018 Feb;212(2):112-118. [doi: 10.1192/bjp.2017.22] [Medline: 29436329]
9. Worm-Smeitink M, Janse A, van Dam A, Evers A, van der Vaart R, Wensing M, et al. Internet-based cognitive behavioral therapy in stepped care for chronic fatigue syndrome: randomized noninferiority trial. J Med Internet Res 2019 Mar 14;21(3):e11276 [FREE Full text] [doi: 10.2196/11276] [Medline: 30869642]
10. Shafran R, Clark DM, Fairburn CG, Arntz A, Barlow DH, Ehlers A, et al. Mind the gap: improving the dissemination of CBT. Behav Res Ther 2009 Nov;47(11):902-909. [doi: 10.1016/j.brat.2009.07.003] [Medline: 19664756]
11. Rothwell PM. External validity of randomised controlled trials: 'to whom do the results of this trial apply?'. Lancet 2005;365(9453):82-93. [doi: 10.1016/S0140-6736(04)17670-8] [Medline: 15639683]
12. Andersson G, Hedman E. Effectiveness of guided internet-based cognitive behavior therapy in regular clinical settings. Verhaltenstherapie 2013;23(3):140-148. [doi: 10.1159/000354779]
13. Vis C, Mol M, Kleiboer A, Bührmann L, Finch T, Smit J, et al. Improving implementation of e-mental health for mood disorders in routine practice: systematic review of barriers and facilitating factors. JMIR Ment Health 2018 Mar 16;5(1):e20 [FREE Full text] [doi: 10.2196/mental.9769] [Medline: 29549072]
14. Vis C, Kleiboer A, Prior R, Bønes E, Cavallo M, Clark SA, et al. Implementing and up-scaling evidence-based e-mental health in Europe: the study protocol for the MasterMind project. Internet Interv 2015 Nov;2(4):399-409. [doi: 10.1016/j.invent.2015.10.002]
15. Crits-Christoph P, Baranackie K, Kurcias J, Beck A, Carroll K, Perry K, et al. Meta‐analysis of therapist effects in psychotherapy outcome studies. Psychother Res 2010 Nov 25;1(2):81-91. [doi: 10.1080/10503309112331335511]
16. Cella M, Stahl D, Reme SE, Chalder T. Therapist effects in routine psychotherapy practice: an account from chronic fatigue syndrome. Psychother Res 2011 Mar;21(2):168-178. [doi: 10.1080/10503307.2010.535571] [Medline: 21271461]
17. Wiborg JF, Knoop H, Wensing M, Bleijenberg G. Therapist effects and the dissemination of cognitive behavior therapy for chronic fatigue syndrome in community-based mental health care. Behav Res Ther 2012 Jun;50(6):393-396. [doi: 10.1016/j.brat.2012.03.002] [Medline: 22504122]
18. Mitchell N, Gordon PK. Attitudes towards computerized CBT for depression amongst a student population. Behav Cogn Psychother 2007 May 14;35(4):421-430. [doi: 10.1017/S1352465807003700]
19. Aerts J, van Dam A. E-health attitudelijst: wat drijft de behandelaar? Psychopraktijk 2015 Oct 27;7(5):26-30. [doi: 10.1007/s13170-015-0090-2]
20. Aerts J, van Dam A. Blended e-health in cognitive behavioural therapy: usage intensity, attitude and therapeutic alliance in clinical practice. Psychol 2018;9(10):2422-2435. [doi: 10.4236/psych.2018.910139]
21. Knoop H, Bleijenberg G. Het Chronisch Vermoeidheidssyndroom: Behandelprotocol Cognitieve Gedragstherapie Voor CVS. Amsterdam, Netherlands: Bohn Stafleu van Loghum; 2010.
22. Janse A, Wiborg JF, Bleijenberg G, Tummers M, Knoop H. The efficacy of guided self-instruction for patients with idiopathic chronic fatigue: a randomized controlled trial. J Consult Clin Psychol 2016 May;84(5):377-388. [doi: 10.1037/ccp0000085] [Medline: 26950098]
23. Vercoulen JH, Swanink CM, Fennis JF, Galama JM, van der Meer JW, Bleijenberg G. Dimensional assessment of chronic fatigue syndrome. J Psychosom Res 1994 Jul;38(5):383-392. [doi: 10.1016/0022-3999(94)90099-X] [Medline: 7965927]
24. Ware Jr JE, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. Med Care 1992 Jun;30(6):473-483. [doi: 10.1097/00005650-199206000-00002] [Medline: 1593914]
25. Wiborg JF, van Bussel J, van Dijk A, Bleijenberg G, Knoop H. Randomised controlled trial of cognitive behaviour therapy delivered in groups of patients with chronic fatigue syndrome. Psychother Psychosom 2015;84(6):368-376. [doi: 10.1159/000438867] [Medline: 26402868]
26. Tummers M, Knoop H, van Dam A, Bleijenberg G. Implementing a minimal intervention for chronic fatigue syndrome in a mental health centre: a randomized controlled trial. Psychol Med 2012 Oct;42(10):2205-2215. [doi: 10.1017/S0033291712000232] [Medline: 22354999]

27. Janse A, Worm-Smeitink M, Bussel-Lagarde J, Bleijenberg G, Nikolaus S, Knoop H. Testing the efficacy of web-based cognitive behavioural therapy for adult patients with chronic fatigue syndrome (CBIT): study protocol for a randomized controlled trial. BMC Neurol 2015 Aug 12;15:137 [FREE Full text] [doi: 10.1186/s12883-015-0392-3] [Medline: 26264735]
28. Worm-Smeitink M, Gielissen M, Bloot L, van Laarhoven HW, van Engelen BG, van Riel P, et al. The assessment of fatigue: psychometric qualities and norms for the checklist individual strength. J Psychosom Res 2017 Jul;98:40-46. [doi: 10.1016/j.jpsychores.2017.05.007] [Medline: 28554371]
29. Stewart AL, Hays RD, Ware Jr JE. The MOS short-form general health survey. Reliability and validity in a patient population. Med Care 1988 Jul;26(7):724-735. [doi: 10.1097/00005650-198807000-00007] [Medline: 3393032]
30. van der Zee KI, Sanderman R. Het Meten Van de Algemene Gezondheidstoestand Met de RAND-36, Een Handleiding. Groningen, Netherlands: Noordelijk Centrum voor Gezondheidsvraagstukken, Rijksuniversiteit Groningen; 2012.
31. Mundt JC, Marks IM, Shear MK, Greist JH. The Work and Social Adjustment Scale: a simple measure of impairment in functioning. Br J Psychiatry 2002 May;180:461-464. [doi: 10.1192/bjp.180.5.461] [Medline: 11983645]
32. Worm-Smeitink M, Nikolaus S, Goldsmith K, Wiborg J, Ali S, Knoop H, et al. Cognitive behaviour therapy for chronic fatigue syndrome: Differences in treatment outcome between a tertiary treatment centre in the United Kingdom and the Netherlands. J Psychosom Res 2016 Aug;87(3):43-49. [doi: 10.1016/j.jpsychores.2016.06.006] [Medline: 27411751]
33. Addis ME, Krasnow AD. A national survey of practicing psychologists' attitudes toward psychotherapy treatment manuals. J Consult Clin Psychol 2000 Apr;68(2):331-339. [doi: 10.1037/0022-006X.68.2.331] [Medline: 10780134]
34. Cohen J. Statistical Power Analysis for the Behavioral Sciences. Second Edition. Hillsdale, New Jersey: Lawrence Erlbaum; 1988.
35. Schmidt FL, Hunter JE. Methods of Meta-Analysis: Correcting Error and Bias in Research Findings. London, UK: Sage Publications; 2014.
36. Jacobson NS, Truax P. Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. J Consult Clin Psychol 1991 Feb;59(1):12-19. [doi: 10.1016/S0005-7894(81)80119-0] [Medline: 2002127]
37. Bower P, Gilbody S. Stepped care in psychological therapies: access, effectiveness and efficiency. Narrative literature review. Br J Psychiatry 2005 Jan;186:11-17. [doi: 10.1192/bjp.186.1.11] [Medline: 15630118]
38. Jaehne A, Loessl B, Frick K, Berner M, Hulse G, Balmford J. The efficacy of stepped care models involving psychosocial treatment of alcohol use disorders and nicotine dependence: a systematic review of the literature. Curr Drug Abuse Rev 2012 Mar;5(1):41-51. [doi: 10.2174/1874473711205010041] [Medline: 22280331]
39. Janse A, van Dam A, Pijpers C, Wiborg JF, Bleijenberg G, Tummers M, et al. Implementation of stepped care for patients with chronic fatigue syndrome in community-based mental health care: outcomes at post-treatment and long-term follow-up. Behav Cogn Psychother 2019 Mar 12:1-11 (epub ahead of print). [doi: 10.1017/S1352465819000110] [Medline: 30859928]
40. Tummers M, Knoop H, van Dam A, Bleijenberg G. Moderators of the treatment response to guided self-instruction for chronic fatigue syndrome. J Psychosom Res 2013 May;74(5):373-377. [doi: 10.1016/j.jpsychores.2013.01.007] [Medline: 23597323]
41. Swift JK, Callahan JL. The impact of client treatment preferences on outcome: a meta-analysis. J Clin Psychol 2009 Apr;65(4):368-381. [doi: 10.1002/jclp.20553] [Medline: 19226606]

Abbreviations

CBT: cognitive behavioral therapy
CDC: Centers for Disease Control and Prevention
CFS: chronic fatigue syndrome
CIS: Checklist Individual Strength
eHealth: electronic health
I-CBT: internet-based cognitive behavioral therapy
ICF: idiopathic chronic fatigue
MHC: mental health care center
RCI: Reliable Change Index
RCT: randomized controlled trial
SF-36: Short Form-36
WSAS: Work and Social Adjustment Scale

Edited by G Eysenbach; submitted 23.03.19; peer-reviewed by H Hadjistavropoulos, A Nakagawa, S Smith; comments to author 25.04.19; revised version received 11.06.19; accepted 30.06.19; published 10.10.19.

Please cite as:
Worm-Smeitink M, van Dam A, van Es S, van der Vaart R, Evers A, Wensing M, Knoop H
Internet-Based Cognitive Behavioral Therapy for Chronic Fatigue Syndrome Integrated in Routine Clinical Care: Implementation Study
J Med Internet Res 2019;21(10):e14037
URL: http://www.jmir.org/2019/10/e14037/
doi:10.2196/14037
PMID:31603428

Original Paper

Estimating the Impact of Novel Digital Therapeutics in Type 2 Diabetes and Hypertension: Health Economic Analysis

Robert J Nordyke[1*], BSc, MSc, PhD; Kevin Appelbaum[2*], BSE; Mark A Berman[2*], BSc, MD

[1]Beta6 Consulting Group, LLC, Topanga, CA, United States

[2]Better Therapeutics, LLC, San Francisco, CA, United States

[*]all authors contributed equally

Corresponding Author:
Mark A Berman, BSc, MD
Better Therapeutics, LLC
445 Bush Street, Suite 300
San Francisco, CA
United States
Phone: 1 617 877 0327
Email: mark@bettertherapeutics.io

Abstract

Background: Behavioral interventions can meaningfully improve cardiometabolic conditions. Digital therapeutics (DTxs) delivering these interventions may provide benefits comparable to pharmacologic therapies, displacing medications for some patients.

Objective: Our objective was to estimate the economic impact of a digital behavioral intervention in type 2 diabetes mellitus (T2DM) and hypertension (HTN) and estimate the impact of clinical inertia on deprescribing medications.

Methods: Decision analytic models estimated health resource savings and cost effectiveness from a US commercial payer perspective. A 3-year time horizon was most relevant to the intervention and payer. Effectiveness of the DTx in improving clinical outcomes was based on cohort studies and published literature. Health resource utilization (HRU), health state utilities, and costs were drawn from the literature with costs adjusted to 2018 dollars. Future costs and quality-adjusted life years (QALYs) were discounted at 3%. Sensitivity analyses assessed uncertainty.

Results: Average HRU savings ranged from $97 to $145 per patient per month, with higher potential benefits in T2DM. Cost-effectiveness acceptability analyses using a willingness-to-pay of $50,000/QALY indicated that the intervention would be cost effective at total 3-year program costs of $6468 and $6620 for T2DM and HTN, respectively. Sensitivity analyses showed that reduced medication costs are a primary driver of potential HRU savings, and the results were robust within values tested. A resistance to deprescribe medications when a patient's clinical outcomes improve can substantially reduce the estimated economic benefits. Our models rely on estimates of clinical effectiveness drawn from limited cohort studies with DTxs and cannot account for other disease management programs that may be implemented. Performance of DTxs in real-world settings is required to further validate their economic benefits.

Conclusions: The DTxs studied may provide substantial cost savings, in part by reducing the use of conventional medications. Clinical inertia may limit the full cost savings of DTxs.

(J Med Internet Res 2019;21(10):e15814) doi:10.2196/15814

KEYWORDS

digital therapeutics; behavioral intervention; economic evaluation; diabetes; hypertension

Introduction

Intensive behavioral and lifestyle interventions have been shown to meaningfully improve clinical outcomes in patients with various cardiometabolic conditions, providing potential for substantial reduction in medication and other resource use. For example, structured, comprehensive lifestyle change programs improve glycemic control in type 2 diabetes mellitus (T2DM), with a substantial number of patients seeing benefits that are comparable or greater than those achieved by pharmacotherapy [1-5]. Behavioral interventions have also demonstrated the ability to control and in some cases achieve normal blood

pressure in patients with hypertension [6,7]. Behavioral interventions have thus established themselves as essential complements to pharmacologic therapy, which is reflected in current treatment guidelines [8-10]. Behavioral interventions also have potential as alternatives to conventional pharmacologic therapy for some patients. For example, participants in the LookAHEAD trial maintained significant improvements over a standard education program in body weight, hemoglobin A_{1c} (HbA_{1c}), systolic blood pressure (SBP), and low-density lipoprotein cholesterol at 4 years, enabling a meaningful proportion of patients to eliminate or reduce antidiabetic pharmacotherapy [1,11].

More recently, mobile software apps have been shown to provide effective platforms to deliver behavioral interventions to patients with cardiometabolic and addictive conditions. The ease of implementation and use of software to treat disease (referred to as a digital therapeutic [DTx]) may help overcome the difficulty of health care systems to deploy intensive behavioral interventions at the large scale needed to improve population outcomes. Moreover, the mobile nature of these apps allows patients to engage with the intervention program several times per day, which may drive improved outcomes over conventional delivery [12]. In cohort studies of T2DM patients, DTxs have demonstrated improvements in clinical outcomes. A mobile medical app that delivered intensive behavioral therapy paired with support from a remote multidisciplinary care team demonstrated mean improvements of HbA_{1c} of 0.8% over a 3-month study period, with improvements up to 1.3% for patients with higher levels of engagement [12]. A similar DTx designed for hypertensive patients demonstrated the ability to reduce blood pressure in a 3-month study with mean reductions in SBP of 11.5 mm Hg and reductions of 17.6 mm Hg among participants with stage 2 hypertension (HTN) [13]. These early results suggest that DTxs may provide clinical benefits comparable to pharmacologic therapy and, in some patients, may help reduce or eliminate the need for medications.

In current clinical practice, however, there is often a delay in deprescribing medications even when the need to do so has been established. This widespread phenomenon, known as clinical inertia, contributes to polypharmacy, which leads to adverse drug reactions, unnecessary costs, and worsened quality of life for patients [14,15]. Clinical inertia, if not addressed, could also lessen the economic benefits realized when a digital therapeutic is put into practice.

The economic benefits of conventionally delivered lifestyle interventions have been demonstrated based on randomized clinical trials [16,17]. However, at this point in their development and introduction into clinical practice, there are few formal evaluations of the potential economic benefits of mobile-platform DTxs and none, that we are aware of, that incorporate measures of clinical inertia. A recent systematic review of a variety of digital health tools showed them to be highly cost effective, although only one study in a Spanish treatment setting evaluated a mobile app comparable to DTxs considered here [18,19]. Thus, our objective in this analysis was to explore the potential economic benefits of DTxs for the treatment of distinct, high-cost cardiometabolic diseases. We developed economic models for the use of DTxs in T2DM and HTN addressing clinical inertia from the perspective of US commercial payers.

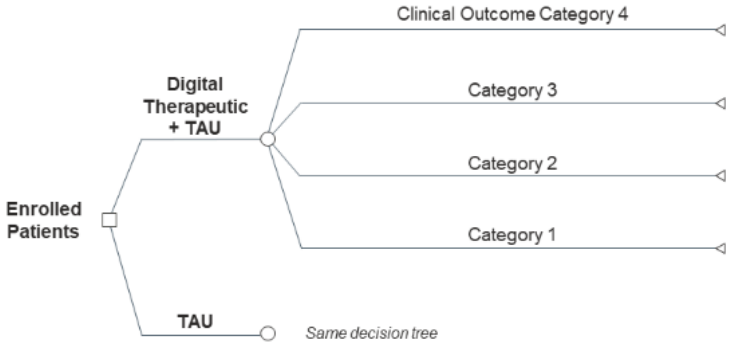
Methods

Model Setting

To understand the relative impact of DTxs in two different cardiometabolic disease states, we created a common framework to estimate the impact of implementing a DTx in distinct patient populations with primary diagnoses of T2DM or HTN. This common framework was implemented in Excel-based decision tree models created from the US payer perspective to best reflect real-world data on medication costs, accurately reflect attrition from the DTx intervention, and account for the currently limited data on DTx effectiveness. A common model framework also facilitates transparent comparisons across disease states modeled. Participants were assumed to enroll in the DTx program at the beginning of year 1 of a 3-year intervention with attrition occurring throughout the program. A 3-year time horizon was chosen as most appropriate for US commercial health plans implementing a behavioral intervention because these plans experience significant annual enrollee turnover and prior studies show that the impact of behavioral intervention can wane over the course of several years [11]. These factors tend to make the longer term benefits to the initiating plan less meaningful financially. We referred to the Consolidated Health Economic Evaluation Reporting Standards (CHEERS) guidelines to improve reporting of this economic analysis [20].

The models were based on the observation that biomarker elevation correlates with both clinical events and health resource utilization (HRU) [21-23]. Each model compared two cohorts (Figure 1): DTx + treatment as usual (TAU) and TAU alone. At this point in their development, economic evaluations of digital health interventions are commonly conducted using a treatment-as-usual comparator [18]. Clinical outcomes were classified into one of four categories chosen to leverage costs reported in the literature and where possible align with current clinical guidelines [8,9,21]. For T2DM, these were defined by HbA_{1c} values of <6.5 (category 1), 6.5 to 7.49% (category 2), 7.5% to 9.0% (category 3), and >9% (category 4). For HTN, these were defined in terms of SBP: <120 mm Hg (category 1), 120 to 129 mm Hg (category 2), 130 to 139 mm Hg (category 3), and >140 mm Hg (category 4). Enrollees were assumed to have active disease with the primary diagnosis corresponding to each model; no patients with already optimized biomarkers were enrolled. Since enrollees were not naïve to conventional pharmacologic treatment, outcomes in the TAU alone group were assumed to be relatively stable over the 3-year time horizon with 80% of TAU alone patients remaining in their outcome category at enrollment.

Figure 1. Model structure.



		Enrolled Patients			
Condition	Parameter	Category 1	Category 2	Category 3	Category 4
T2DM	HbA1c values, %	<6.5	6.5-7.49	7.5-9.0	>9.0
HTN	SBP values, mmHg	<120	120-129	130-139	≥140

Program Attrition

Attrition was considered in the models, including from the DTx and the health plan overall [24]. DTx attrition was considered at several time points. Since long-term attrition rates are not known for DTx, attrition rates were modeled after those commonly seen with pharmacotherapy in HTN and T2DM [25-28]. During year 1, patients were classified as terminating if they did not engage with the app, if they withdrew or clinical outcomes did not improve after 3 months (20% attrition), or if clinical improvements were not durable at year end (additional 20% attrition). Years 2 and 3 also included the withdrawal or lack of durable clinical responses as attrition factors (additional 10% each year). If patients attrited in year 1, their clinical outcomes returned to their enrollment values. During years 2 and 3, program withdrawals were considered to return to average TAU only outcomes for that year. Note that proportions of patients listed as achieving given clinical outcomes at the end of each year (Figure 1 and Multimedia Appendix 1, Figures S1 and S2) are for patients who remained with the DTx + TAU cohort.

Clinical Effectiveness and Clinical Inertia

Model inputs regarding DTx clinical effectiveness are summarized in Table 1. By the end of year 1, the mean change for enrollees remaining in the program was –0.8% (HbA_{1c}) in the T2DM patient population [12] and –11 mm Hg (SBP) in the HTN population [13]. For patients remaining in the DTx program, small improvements were assumed for years 2 and 3 in the base case. Complete descriptions of the 3-year decision trees are provided in Multimedia Appendix 1 (Figures S1 and S2 and Tables S1 and S3).

Benefits from improved clinical outcomes are not assumed to be instantaneous. The base case assumes that for active patients, there were delays of 6 months before medications were reduced based on sustained outcomes and 3 months for any reduction in cardiovascular disease (CVD) hospitalization risk. While we do consider CVD-related hospitalizations, the current evidence base on DTxs does not support accounting for any potential differences in CVD-related mortality. In addition, while many patients in category 2 would be candidates for medication deprescription, we included a parameter controlling the portion of patients in category 2 managed using DTxs alone, without pharmacologic treatment.

Table 1. Model input parameters and ranges for sensitivity analysis.

Parameter	T2DM[a]		HTN[b]	
	Base	Range for SA[c]	Base	Range for SA
Age in years, mean	50	±5	50	±5
Enrolled in category 1, (%)	0	__[d]	0	—
Enrolled in category 2, (%)	0.47	—	0.37	—
Enrolled in category 3, (%)	0.34	—	0.19	—
Enrolled in category 4, (%)	0.19	—	0.44	—
Comorbid conditions, (%)				
T2DM	—	—	33	±20
HC[e]	60	+10/–30	—	—
HTN	60	+10/–30	—	—
Digital therapeutic performance				
Patients improving ≥1 category from baseline, %	62	+10/–33	87	+10/–33
Mean improvement by end of year 1	0.8[f]	+20/–40	11[g]	+20/–40
Medications and resource use				
Category 2 pts not on medications, %	25	0/50	25	0/50
T2DM medications: annual cost ($), range (%)	2466	±20	—	—
T2DM medications: HbA_{1c}[h] gradient for use, slope	0/0.33/1.2/2.2	±10	—	—
HC meds: annual cost ($), range (%)	775	±20	—	—
HC meds: lipid gradient for use, slope	0.5/0.8/1.5/2	±10	—	—
HTN meds: annual cost ($), range (%)	1557	±20	—	—
HTN meds: SBP[i] gradient for use, slope	0/0.15/0.9/1.8	±10	—	—
CVD[j] event cost ($), range (%)	116,423	±20	—	—
HRs[k] of CVD rate by HbA_{1c} level, slope	1/1/1.25/1.98	±10	—	—
Health state utilities (from 0 to 1.0)				
T2DM: category 1 (increment)	0.02	±20	—	—
T2DM: category 2 without medications	0.82	—	—	—
T2DM: category 2 with medications (increment)	–0.02	±20	—	—
T2DM: category 3 (increment)	–0.035	±20	—	—
T2DM: category 4 (increment)	–0.025	±20	—	—
HTN: category 1 (increment)	—	—	0.025	±20
HTN: category 2 without medications	—	—	0.83	—
HTN: category 2 with medications (increment)	—	—	–0.01	±20
HTN: category 3 (increment)	—	—	–0.03	±20
HTN: category 4 (increment)	—	—	0	±20
CVD event (increment)	–0.1	±20	–0.1	±20
Month in year 1 economic benefits realized				
Months required for reduction in medications	6	±3	6	±3
Months required for CVD risk reduction	3	±1	3	±1
Discount rate, %	3	0/5	3	0/5

[a]T2DM: type 2 diabetes mellitus.

[b]HTN: hypertension.
[c]SA: sensitivity analysis.
[d]Not applicable.
[e]HC: high cholesterol.
[f]Hemoglobin A_{1c} level.
[g]mm Hg.
[h]HbA_{1c}: hemoglobin A_{1c}.
[i]SBP: systolic blood pressure.
[j]CVD: cardiovascular disease.
[k]HR: hazard ratio.

Patient Data

Enrolled patients were assumed to be 50% female with a mean age of 50 years in the base case (Table 1). Detailed clinical characteristics by outcome category for each condition were based on LookAHEAD [11], a large prospective study of conventionally delivered intensive lifestyle intervention in T2DM (Multimedia Appendix 1, Table S4). Table 1 also describes the distribution of enrollees by clinical category for each disease state. For T2DM, these are based on data from large US community practices [23]. Distributions of HTN enrollees by clinical category are based on guidelines [9]. Assumed comorbid prevalences are also included in Table 1 [29,30].

Resource Use, Costs, and Health State Utilities

Medication and cardiovascular event costs were based on a survey of the recent literature [31-36]. Medication costs for T2DM [32] do not include insulin costs since we assume the majority of enrolled patients are not insulin dependent. Medication costs for hypertension are estimated for a nationally representative hypertensive patient population [35]. All future costs and benefits were discounted at 3% and adjusted to 2018 dollars.

The models attempt to reflect actual clinical practice by varying medication intensity by disease severity. A recent analysis of administrative claims classified T2DM patients into 4 cohorts based on diabetes-related drug utilization [21]. The study found that patients with HbA_{1c} >8.9% had diabetes-related drug costs over 9 times those who recently initiated monotherapy with an average HbA_{1c} of 8.0%. Patients moderately controlled on monotherapy were 1.43 times more costly than diagnosed patients without treatment (mean baseline HbA_{1c} of 6.4%), while those poorly controlled were 2.44 to 2.98 times more costly. An analysis of commercial health plan data examined resource utilization among adult T2DM patients categorized by HbA_{1c} at baseline [37]. With HbA_{1c} <7.0% as the reference, patients with HbA_{1c} levels of 7.0% to 8.0% had 1-year prescription costs 45% higher, those with levels 8.0% to 9.0% were 108% higher, and those with levels >9.0% were 131% higher. Similar relationships between HbA_{1c} and diabetes-related hospitalizations were found in an analysis of 200,000 patients with either type 1 or type 2 diabetes [38].

Relationships between medication intensity and disease severity have not been as thoroughly studied in HTN, although these trends are clearly reflected in treatment guidelines. Medication intensity gradients assumed for HTN are weaker than the literature-based estimate used in the T2DM model. CVD event rates are based on hazard ratios by HbA_{1c} level [22] and by the Framingham 10-Year Risk of General Cardiovascular Disease risk equation for HTN-specific outcomes [39].

Health state utilities were drawn from the literature [40,41]. Baseline utilities are defined for category 2, with increments for improved status or less drug utilization and utility decrements for worse health states (see Multimedia Appendix 1 tables for calculated utilities for all health states in the models).

Sensitivity Analyses

One-way deterministic sensitivity analyses were performed to assess robustness of the base case evaluation of potential health resource saving. Most parameters varied by ±20% (Table 1) with the primary exception being DTx effectiveness. Since real-world effectiveness of any single intervention may be lower than demonstrated in a controlled small-cohort setting, our sensitivity analyses assume a conservative, asymmetric range of 40% below the base case estimate and only 20% above.

Results

Savings in T2DM are estimated at $83 per participant per month (PPPM) in year 1 and rise to $174 to $178 in years 2 and 3, respectively. Year 1 savings in HTN are estimated at $70 PPPM rising to $113 and $107 in years 2 and 3, respectively. Estimated year 1 savings in HRU (Figure 2) are lower than savings in subsequent years due to delays in realizing economic benefits of reducing medications and CVD-related hospitalizations. The primary driver of these differences across disease states is the magnitude of potential reduction in medication costs, with higher average drug costs in T2DM. In addition, the Framingham equation reflects a less steep CVD risk trend across the outcome categories in HTN compared with the trend between CVD risk and HbA_{1c} [22].

The estimated savings reflect changes in clinical effectiveness over time (Figure 1 and Multimedia Appendix 1, Figures S1 and S2). The largest improvements in outcomes are assumed in year 1, with incremental improvements in years 2 and 3 for DTx + TAU. The most important factor of clinical effectiveness driving estimated economic benefits is the differential between the two cohorts in each patient population (Multimedia Appendix 1, Table S4). The effectiveness of TAU alone is comparable in both patient populations, accounting for differences in severity at enrollment. However, DTx effectiveness in HTN is assumed to be slightly more effective

than in T2DM. The majority of estimated HRU savings are due to potential reductions in medication costs for both patient populations and across program years. Medication reductions contribute a larger proportion of savings in years 2 and 3 for T2DM due to higher assumed average medication costs and clinical inertia (Figure 2). Due to the shallow trend in CVD risk across HTN outcome categories, inpatient costs represent a relatively smaller contribution to total savings.

In the sensitivity analyses (Figure 3), assumptions regarding the distribution of enrollee disease severity are a significant driver of uncertainty for the T2DM population, although this is less important in HTN. In the high-cost, high-effectiveness scenario, only more severe patients (categories 3 and 4) are enrolled (0% category 2). Since the base case assumes nearly half of enrolled T2DM patients are in category 2, estimated savings rise when more patients can experience larger improvements in clinical outcomes. Notably, assumed HRU costs are not the largest driver of uncertainty in year 1, while medication and hospitalization costs are an important driver of uncertainty in subsequent years. Notably, delays in realizing economic benefits of improving HbA_{1c} and SBP levels are a large driver of uncertainty for year 1 and moderately important in years 2 and 3. In T2DM, for example, greater delays in realizing benefits reduces estimated HRU savings to $55 PPPM from the baseline of $83, while shorter delays (3 months vs the baseline 6 months) result in estimated savings of $98 PPPM in year 1. Clinical inertia is a smaller contributor to uncertainty in years 2 and 3 for T2DM, resulting in variances of about 15% for those years. Estimates for HTN indicate that clinical inertia assumptions vary HRU savings by 42% to 47% in year 1. Clinical inertia becomes a small driver of uncertainty in subsequent years in HTN. Assumptions regarding DTx clinical effectiveness (Figure 1 and Multimedia Appendix 1, Figures S1 and S2) are important drivers of uncertainty. However, in the T2DM population, these assumptions are less important than HRU costs, severity distribution, clinical inertia, and comorbidities. Also, with the lower drug costs and weaker trend in CVD risk in HTN, clinical effectiveness is a relatively greater driver in this condition.

The threshold analysis (Figure 4) examines cost effectiveness with varying levels of total DTx program costs over the 3-year time horizon. At a willingness-to-pay threshold of $100,000/quality-adjusted life year (QALY), the DTx + TAU combination is estimated to be cost effective at total 3-year program costs of $8348 (T2DM) and $10,212 (HTN). At threshold of $50,000/QALY, these values are $6468 (T2DM) and $6620 (HTN). These estimates are less elastic for T2DM due to higher drug costs and stronger relationships between outcome categories and CVD hospitalization.

Figure 2. Base case health resource use savings and contributions to estimated savings. Cost estimates are per enrollee per month in year 1 dollars by patient population.

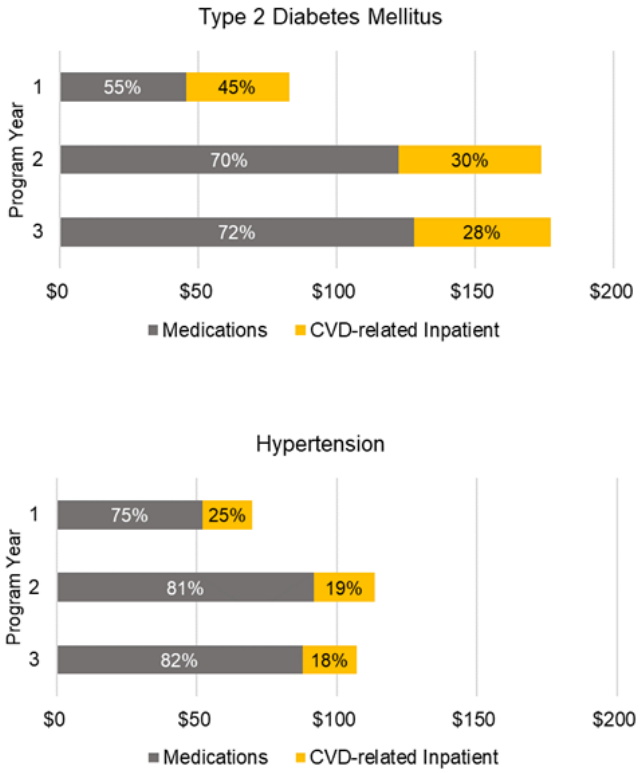


Figure 3. Health resource use sensitivity analysis by patient population.

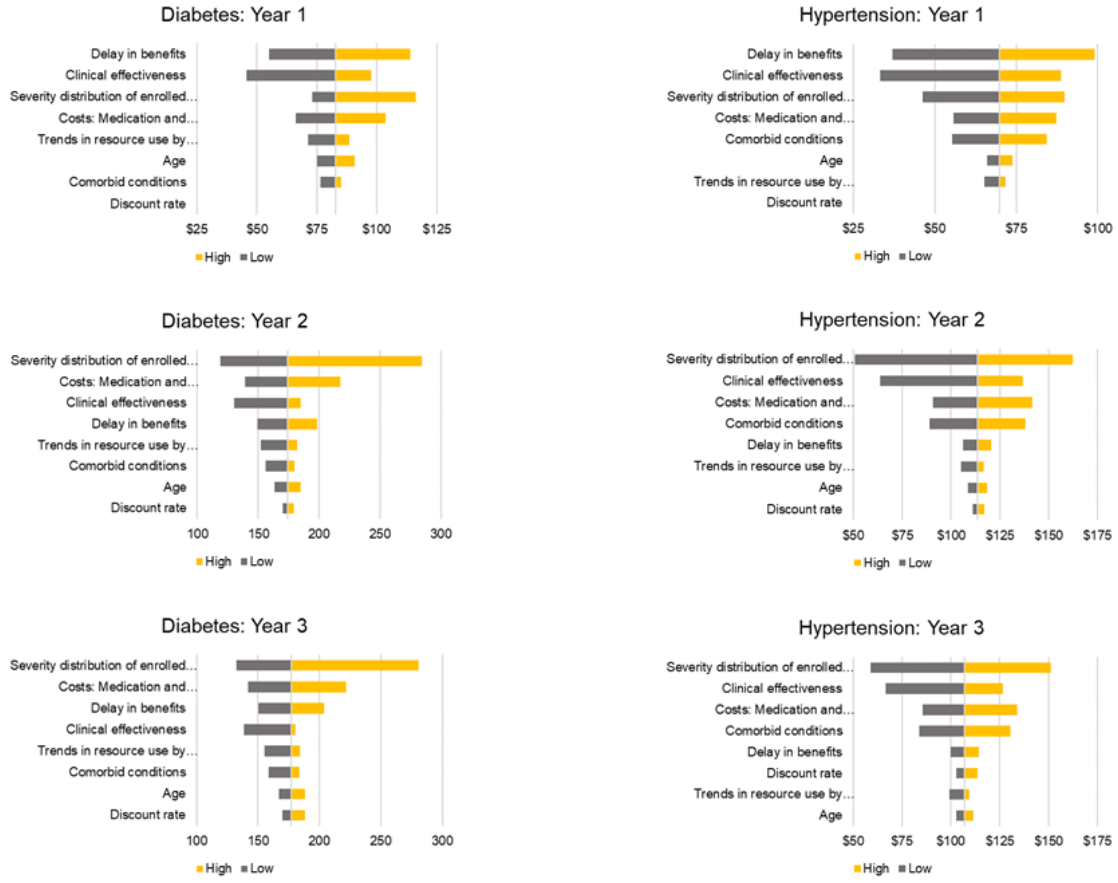
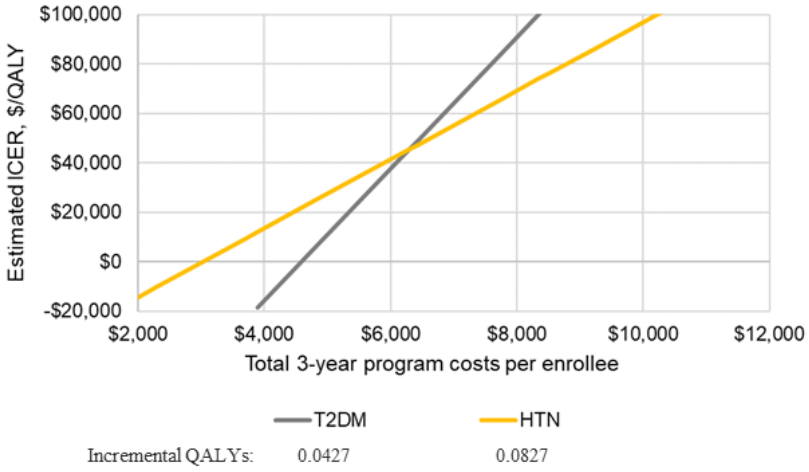


Figure 4. Cost-effectiveness threshold curves by patient population.



Discussion

Principal Findings

The addition of DTxs to conventional pharmacologic treatment as usual in cardiometabolic diseases holds the potential to reduce HRU costs. Sensitivity analyses show that potential HRU savings were sensitive to assumptions regarding the magnitude of HRU costs offset by the DTx, severity distribution of enrolled patients, estimates of DTx clinical effectiveness, and measures of clinical inertia. Cost-effectiveness analyses, limited by the 3-year time horizon, estimated that at a willingness-to-pay threshold of only $50,000/QALY, addition of the DTx would be cost effective at total 3-year DTx intervention costs of up to the average total medication costs for these diseases over the same period.

This study has demonstrated some differential impacts of DTxs in two cardiometabolic diseases and suggests hypotheses for further exploration. One area is the finding that the severity of

enrolled patients greatly affects the potential benefits of the DTx. Sensitivity analyses (Figure 3) showed that severity at enrollment was the third largest driver of uncertainty in year 1 and the largest in years 2 and 3. Restricting enrollment to a moderately severe to severe population (categories 3 and 4 only) increased estimated PPPM HRU savings by 40% to 60% in T2DM and 30% to 40% in HTN. As an extension, we also examined the impact (results not shown) of enrolling only the most severe (category 4) patients, which resulted in lower HRU savings in years 2 and 3 relative to enrolling category 3 and 4 patients. This is due to our modeled assumptions that enrolled category 4 patients are more resistant to improvements than are category 3 patients. The impact of severity at enrollment varies across disease states and is due in part to the distribution of HTN versus T2DM severity in a typical commercially insured population. For example, there are far more severe (category 4) enrolled HTN patients than category 4 T2DM patients, and there is a small proportion of moderately severe (category 3) enrolled HTN patients. Estimated gains in both T2DM and HTN are substantial. However, most improvements occur in year 1 for HTN patients with modest improvements in years 2 and 3. Whereas in T2DM, with fewer severe (category 4) patients enrolled, the estimated savings in years 2 and 3 are over twice those estimated in year 1. Future economic evaluations or projections of DTxs should consider the effect of baseline disease severity on the cost benefits of the treatment.

Another hypothesis relates to how clinical inertia could limit the economic benefits realized from digitally delivered behavioral interventions that treat cardiometabolic diseases. As an example, participants who experience a sustained improvement in HbA_{1c} below 6.5% should be considered for medication reduction. Research is needed to quantify the extent of clinical inertia observed in real-world implementations of DTxs. In addition, apps that work with DTxs, such as clinical decision support software, should be explored as solutions to clinical inertia.

Limitations

Our analysis has several limitations. First, our results are based on simple decision analytic models that rely on estimates of clinical effectiveness drawn from limited cohort studies with DTxs. T2DM and HTN are complex, chronic conditions, and sophisticated techniques coupled with detailed effectiveness data are required to accurately simulate treatment outcomes over a 10- to 20-year time period. However, over the short time horizon appropriate in this setting, our approach likely provides valid directional estimates of potential benefits to US commercial payers. In addition, since patient-level benefits will likely continue after payer reimbursements end, the relatively short time horizon will generate conservative estimates of cost effectiveness. While the potential improvements in clinical outcomes for patients responding to DTxs are consistent with prior experience (eg, an average 0.8% reduction in HbA_{1c} for T2DM patients at 13 weeks), sensitivity analyses confirmed that these are important drivers of potential savings. Performance of DTxs in real-world settings is required to further validate their potential for cost savings. An important driver of real-world effectiveness is program attrition. While we account for attrition throughout the program, with 36% of enrollees withdrawing by the end of year 1 alone, attrition in actual practice may be higher, reducing the total economic savings possible. Second, the analyses assume that the DTx is the only intervention alongside treatment as usual. This is relatively common in economic evaluations of digital health [18]. However, in practice, particularly in US commercial health plans, patients may participate in multiple cardiometabolic disease management programs. In such situations, it will be difficult to associate the specific impact of one intervention versus another, and the net result is uncertain a priori. To the extent that there is overlap in content of a DTx and conventional interventions, the improvements attributable solely to the DTx may be less than estimated in our modeled scenarios. Conversely, there may be synergistic effects between DTxs and conventional interventions, with larger net benefits. More research is required on the best ways to implement DTxs, including analyses of real-world observational data of DTxs using econometric or machine learning methods to distinguish individual effects of multiple interventions. Additionally, no costs of adverse events were included in these analyses. For patients with the same clinical outcome category receiving similar medication regimens, the medication-related adverse events and associated costs would be comparable across DTx + TAU versus TAU alone cohorts. However, no known clinical adverse events are associated with the DTx, and a larger proportion of patients will be managed to improved outcomes not requiring pharmacologic treatment while using the DTx. Thus, not including adverse events may bias the estimates of DTx benefits downward slightly. Finally, while we account for attrition from the DTx cohort, we don't directly take treatment adherence to TAU into account. However, our drug cost estimates are drawn from published analyses of commercial claims data for diagnosed patients, which account for some level of nonadherence in a typical health plan.

Comparison With Prior Work

Averaged over the 3-year time horizon, estimated savings were $145 PPPM for the T2DM population and $97 PPPM in HTN patients. Given the drug costs for T2DM and HTN drawn from the peer-reviewed literature, these savings represent approximately 22% to 29% of the total estimated medical costs for an average patient treated as usual. This estimate is in line with the findings of the economics benefits of a mobile DTx for heart failure patients, which found a 33% reduction in total management and treatment costs, and an app-based glucose monitoring program, which found a 22% reduction in total medical spending [19,42].

Conclusions

DTXs for T2DM and HTN patients may provide substantial improvements in patient outcomes resulting in lower HRU and costs when compared with standard pharmacologic-based treatment as usual. Clinical inertia may be a barrier to realizing the benefits of DTxs.

Acknowledgments

This study was funded by Better Therapeutics, LLC. The authors would like to thank the entire team at Better Therapeutics for their dedication to building and evaluating digital therapeutics.

Authors' Contributions

All authors made substantial contributions to the design of the research and interpretation of results. All authors also either drafted the manuscript or revised it critically for important content, and all authors provided final approval of the version to be published.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary figures and tables.
[PDF File (Adobe PDF File), 393 KB - jmir_v21i10e15814_app1.pdf]

References

1. Gregg EW, Chen H, Wagenknecht LE, Clark JM, Delahanty LM, Bantle J, Look AHEAD Research Group. Association of an intensive lifestyle intervention with remission of type 2 diabetes. JAMA 2012 Dec 19;308(23):2489-2496 [FREE Full text] [doi: 10.1001/jama.2012.67929] [Medline: 23288372]
2. Barnard ND, Cohen J, Jenkins DJA, Turner-McGrievy G, Gloede L, Green A, et al. A low-fat vegan diet and a conventional diabetes diet in the treatment of type 2 diabetes: a randomized, controlled, 74-wk clinical trial. Am J Clin Nutr 2009 May;89(5):1588S-1596S [FREE Full text] [doi: 10.3945/ajcn.2009.26736H] [Medline: 19339401]
3. Colberg SR, Sigal RJ, Fernhall B, Regensteiner JG, Blissmer BJ, Rubin RR, et al. Exercise and type 2 diabetes: the American College of Sports Medicine and the American Diabetes Association: joint position statement executive summary. Diabetes Care 2010 Dec;33(12):2692-2696 [FREE Full text] [doi: 10.2337/dc10-1548] [Medline: 21115771]
4. Chen L, Pei J, Kuang J, Chen H, Chen Z, Li Z, et al. Effect of lifestyle intervention in patients with type 2 diabetes: a meta-analysis. Metabolism 2015 Feb;64(2):338-347 [FREE Full text] [doi: 10.1016/j.metabol.2014.10.018] [Medline: 25467842]
5. Johansen MY, MacDonald CS, Hansen KB, Karstoft K, Christensen R, Pedersen M, et al. Effect of an intensive lifestyle intervention on glycemic control in patients with type 2 diabetes: a randomized clinical trial. JAMA 2017 Dec 15;318(7):637-646. [doi: 10.1001/jama.2017.10169] [Medline: 28810024]
6. Homan SG, McBride DG, Yun S. The effect of the Missouri WISEWOMAN program on control of hypertension, hypercholesterolemia, and elevated blood glucose among low-income women. Prev Chronic Dis 2014 May 01;11:E74 [FREE Full text] [doi: 10.5888/pcd11.130338] [Medline: 24784910]
7. Patnode C, Evans C, Senger C, Redmond N, Lin J. Behavioral counseling to promote a healthful diet and physical activity for cardiovascular disease prevention in adults without known cardiovascular disease risk factors: updated systematic review for the U.S. Preventive Services Task Force. Rockville: Agency for Healthcare Research and Quality; 2017. URL: https://www.ncbi.nlm.nih.gov/books/n/es152/pdf/ [accessed 2019-09-25]
8. Association American Diabetes. Updates to “Standards of Medical Care in Diabetes–2018”. Diabetes Care 2018 Sep;41(9):2045-2047. [doi: 10.2337/dc18-su09] [Medline: 30135199]
9. Whelton PK, Carey RM, Aronow WS, Casey DE, Collins KJ, Dennison Himmelfarb C, et al. 2017 ACC/AHA/AAPA/ABC/ACPM/AGS/APhA/ASH/ASPC/NMA/PCNA Guideline for the prevention, detection, evaluation, and management of high blood pressure in adults: executive summary: a report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. Circulation 2018 Oct 23;138(17):e426-e483. [doi: 10.1161/CIR.0000000000000597] [Medline: 30354655]
10. Jellinger PS, Handelsman Y, Rosenblit PD, Bloomgarden ZT, Fonseca VA, Garber AJ, et al. American Association of Clinical Endocrinologists and American College of Endocrinology Guidelines for Management of Dyslipidemia And Prevention of Cardiovascular Disease. Endocr Pract 2017 Apr;23(Suppl 2):479-497. [doi: 10.4158/EP171764.APPGL] [Medline: 28437620]
11. Look AHEAD Research Group, Wing RR. Long-term effects of a lifestyle intervention on weight and cardiovascular risk factors in individuals with type 2 diabetes mellitus: four-year results of the Look AHEAD trial. Arch Intern Med 2010 Sep 27;170(17):1566-1575 [FREE Full text] [doi: 10.1001/archinternmed.2010.334] [Medline: 20876408]
12. Berman MA, Guthrie NL, Edwards KL, Appelbaum KJ, Njike VY, Eisenberg DM, et al. Change in glycemic control with use of a digital therapeutic in adults with type 2 diabetes: cohort study. JMIR Diabetes 2018 Feb 14;3(1):e4 [FREE Full text] [doi: 10.2196/diabetes.9591] [Medline: 30291074]

13. Guthrie N, Berman M, Edwards K, Appelbaum K, Dey S, Carpenter J, et al. Achieving rapid blood pressure control with digital therapeutics: retrospective cohort and machine learning study. JMIR Cardio 2019 Mar 12;3(1):e13030. [doi: 10.2196/13030]
14. Reach G, Pechtner V, Gentilella R, Corcos A, Ceriello A. Clinical inertia and its impact on treatment intensification in people with type 2 diabetes mellitus. Diabetes Metab 2017 Dec;43(6):501-511 [FREE Full text] [doi: 10.1016/j.diabet.2017.06.003] [Medline: 28754263]
15. Reeve E, Thompson W, Farrell B. Deprescribing: a narrative review of the evidence and practical recommendations for recognizing opportunities and taking action. Eur J Intern Med 2017 Mar;38:3-11. [doi: 10.1016/j.ejim.2016.12.021] [Medline: 28063660]
16. Peels DA, Hoogenveen RR, Feenstra TL, Golsteijn RH, Bolman C, Mudde AN, et al. Long-term health outcomes and cost-effectiveness of a computer-tailored physical activity intervention among people aged over fifty: modelling the results of a randomized controlled trial. BMC Public Health 2014 Oct 23;14:1099 [FREE Full text] [doi: 10.1186/1471-2458-14-1099] [Medline: 25342517]
17. Jacobs-van der Bruggen MAM, van Baal PH, Hoogenveen RT, Feenstra TL, Briggs AH, Lawson K, et al. Cost-effectiveness of lifestyle modification in diabetic patients. Diabetes Care 2009 Aug;32(8):1453-1458 [FREE Full text] [doi: 10.2337/dc09-0363] [Medline: 19435958]
18. Jiang X, Ming W, You JH. The cost-effectiveness of digital health interventions on the management of cardiovascular diseases: systematic review. J Med Internet Res 2019 Jun 17;21(6):e13166 [FREE Full text] [doi: 10.2196/13166] [Medline: 31210136]
19. Cano Martín JA, Martínez-Pérez B, de la Torre-Díez I, López-Coronado M. Economic impact assessment from the use of a mobile app for the self-management of heart diseases by patients with heart failure in a Spanish region. J Med Syst 2014 Sep;38(9):96. [doi: 10.1007/s10916-014-0096-z] [Medline: 24994514]
20. Husereau D, Drummond M, Petrou S, Carswell C, Moher D, Greenberg D, et al. Consolidated Health Economic Evaluation Reporting Standards (CHEERS)—explanation and elaboration: a report of the ISPOR Health Economic Evaluation Publication Guidelines Good Reporting Practices Task Force. Value Health 2013;16(2):231-250 [FREE Full text] [doi: 10.1016/j.jval.2013.02.002] [Medline: 23538175]
21. Bonafede M, Chandran A, DiMario S, Saltiel-Berzin R, Saliu D. Medication usage, treatment intensification, and medical cost in patients with type 2 diabetes: a retrospective database study. BMJ Open Diabetes Res Care 2016;4(1):e000189 [FREE Full text] [doi: 10.1136/bmjdrc-2015-000189] [Medline: 27547410]
22. Nichols GA, Joshua-Gotlib S, Parasuraman S. Glycemic control and risk of cardiovascular disease hospitalization and all-cause mortality. J Am Coll Cardiol 2013 Jul 09;62(2):121-127 [FREE Full text] [doi: 10.1016/j.jacc.2013.04.031] [Medline: 23665365]
23. Wilson PWF. Lipids and vascular disease: a Framingham perspective. Glob Heart 2013 Mar;8(1):25-33 [FREE Full text] [doi: 10.1016/j.gheart.2012.12.009] [Medline: 25690261]
24. Cunningham PJ, Kohn L. Health plan switching: choice or circumstance? Health Aff (Millwood) 2000;19(3):158-164. [Medline: 10812794]
25. Durand H, Hayes P, Morrissey EC, Newell J, Casey M, Murphy AW, et al. Medication adherence among patients with apparent treatment-resistant hypertension: systematic review and meta-analysis. J Hypertens 2017 Dec;35(12):2346-2357. [doi: 10.1097/HJH.0000000000001502] [Medline: 28777133]
26. Krass I, Schieback P, Dhippayom T. Adherence to diabetes medication: a systematic review. Diabet Med 2015 Jun;32(6):725-737. [doi: 10.1111/dme.12651] [Medline: 25440507]
27. Lee WC, Balu S, Cobden D, Joshi AV, Pashos CL. Prevalence and economic consequences of medication adherence in diabetes: a systematic literature review. Manag Care Interface 2006 Jul;19(7):31-41. [Medline: 16898343]
28. Abegaz TM, Shehab A, Gebreyohannes EA, Bhagavathula AS, Elnour AA. Nonadherence to antihypertensive drugs: a systematic review and meta-analysis. Medicine (Baltimore) 2017 Jan;96(4):e5641 [FREE Full text] [doi: 10.1097/MD.0000000000005641] [Medline: 28121920]
29. Courtemanche T, Mansueto G, Hodach R, Handmaker K. Population health approach for diabetic patients with poor A1C control. Am J Manag Care 2013 Jun;19(6):465-472 [FREE Full text] [Medline: 23844708]
30. DeVore AD, Sorrentino M, Arnsdorf MF, Ward RP, Bakris GL, Blankstein R. Predictors of hypertension control in a diverse general cardiology practice. J Clin Hypertens (Greenwich) 2010 Aug;12(8):570-577 [FREE Full text] [doi: 10.1111/j.1751-7176.2010.00298.x] [Medline: 20695933]
31. Levy PD, Willock RJ, Burla M, Brody A, Mahn J, Marinica A, et al. Total antihypertensive therapeutic intensity score and its relationship to blood pressure reduction. J Am Soc Hypertens 2016 Dec;10(12):906-916 [FREE Full text] [doi: 10.1016/j.jash.2016.10.005] [Medline: 27856202]
32. American Diabetes Association. Economic costs of diabetes in the U.S. in 2012. Diabetes Care 2013 Apr;36(4):1033-1046 [FREE Full text] [doi: 10.2337/dc12-2625] [Medline: 23468086]
33. Dehmer SP, Maciosek MV, LaFrance AB, Flottemesch TJ. Health benefits and cost-effectiveness of asymptomatic screening for hypertension and high cholesterol and aspirin counseling for primary prevention. Ann Fam Med 2017 Jan;15(1):23-36 [FREE Full text] [doi: 10.1370/afm.2015] [Medline: 28376458]

34. Moran AE, Odden MC, Thanataveerat A, Tzong KY, Rasmussen PW, Guzman D, et al. Cost-effectiveness of hypertension therapy according to 2014 guidelines. N Engl J Med 2015 Jan 29;372(5):447-455 [FREE Full text] [doi: 10.1056/NEJMsa1406751] [Medline: 25629742]
35. Kirkland EB, Heincelman M, Bishu KG, Schumann SO, Schreiner A, Axon RN, et al. Trends in healthcare expenditures among US adults with hypertension: national estimates, 2003-2014. J Am Heart Assoc 2018 May 30;7(11):e008731 [FREE Full text] [doi: 10.1161/JAHA.118.008731] [Medline: 29848493]
36. Fitch K, Iwasaki K, Pyenson B. The cost and quality gap in diabetes care: an actuarial analysis. 2012 Jan 30. URL: http://www.milliman.com/uploadedFiles/insight/health-published/cost-quality-gap-diabetes.pdf [accessed 2019-09-25]
37. Meng J, Casciano R, Lee Y, Stern L, Gultyaev D, Tong L, et al. Effect of diabetes treatment-related attributes on costs to type 2 diabetes patients in a real-world population. J Manag Care Spec Pharm 2017 Apr;23(4):446-452 [FREE Full text] [doi: 10.18553/jmcp.2017.23.4.446] [Medline: 28345434]
38. Menzin J, Korn JR, Cohen J, Lobo F, Zhang B, Friedman M, et al. Relationship between glycemic control and diabetes-related hospital costs in patients with type 1 or type 2 diabetes mellitus. J Manag Care Pharm 2010 May;16(4):264-275 [FREE Full text] [doi: 10.18553/jmcp.2010.16.4.264] [Medline: 20433217]
39. D'Agostino RB, Vasan RS, Pencina MJ, Wolf PA, Cobain M, Massaro JM, et al. General cardiovascular risk profile for use in primary care: the Framingham Heart Study. Circulation 2008 Feb 12;117(6):743-753 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.107.699579] [Medline: 18212285]
40. Beaudet A, Clegg J, Thuresson P, Lloyd A, McEwan P. Review of utility values for economic modeling in type 2 diabetes. Value Health 2014 Jun;17(4):462-470 [FREE Full text] [doi: 10.1016/j.jval.2014.03.003] [Medline: 24969008]
41. Zhang P, Brown MB, Bilik D, Ackermann RT, Li R, Herman WH. Health utility scores for people with type 2 diabetes in U.S. managed care health plans: results from Translating Research Into Action for Diabetes (TRIAD). Diabetes Care 2012 Nov;35(11):2250-2256 [FREE Full text] [doi: 10.2337/dc11-2478] [Medline: 22837369]
42. Whaley CM, Bollyky JB, Lu W, Painter S, Schneider J, Zhao Z, et al. Reduced medical spending associated with increased use of a remote diabetes management program and lower mean blood glucose values. J Med Econ 2019 May 08:1-9. [doi: 10.1080/13696998.2019.1609483] [Medline: 31012392]

Abbreviations

CHEERS: Consolidated Health Economic Evaluation Reporting Standards
CVD: cardiovascular disease
DTx: digital therapeutic
HbA_{1c}: hemoglobin A_{1c}
HRU: health resource utilization
HTN: hypertension
PPPM: per participant per month
QALY: quality-adjusted life year
SBP: systolic blood pressure
T2DM: type 2 diabetes mellitus
TAU: treatment as usual

Edited by G Eysenbach; submitted 08.08.19; peer-reviewed by A Davis, WK Ming; comments to author 10.09.19; revised version received 18.09.19; accepted 20.09.19; published 09.10.19.

Please cite as:
Nordyke RJ, Appelbaum K, Berman MA
Estimating the Impact of Novel Digital Therapeutics in Type 2 Diabetes and Hypertension: Health Economic Analysis
J Med Internet Res 2019;21(10):e15814
URL: https://www.jmir.org/2019/10/e15814
doi:10.2196/15814
PMID:31599740

Original Paper

Process Evaluation of Nurse-Led Online Self-Management Support for Family Caregivers to Deal With Behavior Changes of a Relative With Dementia (Part 1): Mixed Methods Study

Judith G Huis in het Veld[1], MSc; Iris F M van Asch[2], MSc; Bernadette M Willemse[2], MSc, PhD; Paul-Jeroen Verkade[3], MSc; Anne Margriet Pot[4], MSc, PhD; Marco M Blom[5], MSc; Rob B M Groot Zwaaftink[5], MSc; Anneke L Francke[1,6], MSc, PhD

[1]Amsterdam UMC, Vrije Universiteit Amsterdam, Department of Public and Occupational Health, Amsterdam Public Health Research Institute, Amsterdam, Netherlands

[2]Netherlands Institute of Mental Health and Addiction, Trimbos Institute, Utrecht, Netherlands

[3]The Geriant Foundation, Region North of Amsterdam, Netherlands

[4]World Health Organization, Geneva, Switzerland

[5]Dutch Alzheimer's Society, Amersfoort, Netherlands

[6]Netherlands Institute for Health Services Research (NIVEL), Utrecht, Netherlands

Corresponding Author:
Judith G Huis in het Veld, MSc
Amsterdam UMC, Vrije Universiteit Amsterdam, Department of Public and Occupational Health, Amsterdam Public Health Research Institute
Van der Boechorststraat 7
Amsterdam, 1081 BT
Netherlands
Phone: 31 20 44 45 365
Email: j.huisinhetveld@vumc.nl

Abstract

Background: Coping with behavioral changes is a daily challenge for family caregivers in all phases of dementia, and assistance is needed for it. An online self-management support intervention was therefore developed and conducted involving the following elements: (1) email contact with a specialized dementia nurse, (2) online videos, and (3) e-bulletins containing information about behavior changes and how to manage them.

Objective: The aim of this study was to understand (1) family caregivers' actual use of various elements of the online self-management support, (2) family caregivers' evaluation and satisfaction with the various elements, and (3) nurses' usage and evaluations of the online support through the tailored email contacts.

Methods: A mixed methods design was used in this process evaluation, combining quantitative and qualitative methods including analyses of dementia nurses' registration forms, the number of clicks on online videos and e-bulletins, evaluation questions answered by family caregivers in a survey questionnaire, semistructured interviews with family caregivers and nurses, and analysis of the content of the email contacts.

Results: The actual use of various elements of the online self-management support by family caregivers varied: 78% (21/27) of family caregivers had an email contact with the specialist nurse, 80% (43/54) of family caregivers clicked on an online video, and 37% (30/81) clicked on an e-bulletin. Family caregivers showed positive evaluations and satisfaction. The tailor-made approach in the personal email contacts in particular was valued by the family caregivers. Nurses' evaluations about providing self-management support online were mixed as it was a relatively new task for them.

Conclusions: An important insight is that not all participants made optimum use of the various elements of the intervention. Nurses also said that the email contacts were more often used to express feelings about coping with behavioral changes. More research is needed to investigate the reasons why people accept, adopt, and adhere to online interventions to reduce cases where they are not used and to back them up appropriately with tailored (online) information and advice for their personal situations.

(J Med Internet Res 2019;21(10):e13002) doi:10.2196/13002

KEYWORDS

dementia; internet; eHealth; caregiver

Introduction

Background

Family caregivers of people with dementia often face many challenges in everyday life while caring for their relative [1], most prominently regarding changes in behavior of the person with dementia [2,3]. People with dementia may exhibit behavior that is dependent, aggressive, suspicious, apathetic, or indifferent, or night-time restlessness and masking behavior. Approximately 80% to 90% of people with dementia show behavior disturbances during the disease process [4], often distressing both the person with dementia and their family caregivers [3,5].

Coping with behavioral changes is a daily challenge for family caregivers in all phases of dementia [6]. These days, the term *self-management* is widely used when referring to managing consequences of a disease in daily life. Barlow et al [7] defined self-management as *the individual's ability to manage the symptoms, treatment, physical and psychosocial consequences and lifestyle changes inherent in living with a chronic condition.* Self-management applies not only to the patient but also to family caregivers. Especially in dementia care, the person often becomes increasingly dependent on support from family caregivers. This is stressful for family caregivers, especially when coping with behavioral changes [5,6]. They use strategies to respond to behavioral changes by remaining calm or encouraging activities and distractions. Moreover, family caregivers have self-management strategies to manage their own caregiver stress and problems related to their relative's dementia [5,8].

However, family caregivers might need assistance coping with this daily challenge. In particular, nurses are in the best position to help them because they develop a close partnership with individuals and their families throughout their lives [9]. This nurse-patient contact can also occur online [10] and might be especially useful for reaching family caregivers who are short of time because of providing care, have transportation difficulties or do not want to leave the person with dementia alone at home [11,12]. In addition to professional support online, family caregivers may also benefit from multicomponent online interventions that combine, for example, information and tailored caregiving strategies [13].

In this paper, we present a process evaluation of an online self-management support intervention addressing behavioral changes in dementia. The intervention consists of various online elements. The process evaluation was performed alongside a randomized controlled trial (RCT) [14]. The aim of the RCT was to explore (1) whether a major online self-management support intervention involving email contacts with a specialist dementia nurse in combination with online videos and e-bulletins is more effective than minor interventions not involving the email contacts with the nurse, and (2) if a medium intervention including videos and e-bulletins is more effective than a minor intervention including e-bulletins only. The results showed no statistically significant effects on family caregivers' self-efficacy for the major and medium online self-management support interventions compared with the minor intervention [15].

It is important to carry out a process evaluation alongside RCTs to allow effects (or the lack thereof) to be interpreted. It enables researchers to understand whether and how interventions are used, and how interventions are being evaluated by the people involved. Process evaluations alongside RCTs are even more important when evaluating online interventions because these studies are complicated, given the high numbers of nonadherent participants compared with face-to-face interventions [16,17].

Objective

The overall objective of the process evaluation was to get an idea of the actual usage and evaluations of the intervention components. Related subobjectives were to understand (1) actual usage by family caregivers of the various elements of the online support, (2) family caregivers' evaluation of and satisfaction with the various elements, and (3) nurses' usage and evaluations of the online support through the tailored email contacts.

Methods

Design

The process evaluation had a mixed methods design in which quantitative and qualitative methods were combined, and various sources were used (see the *Data Collection* section). The process evaluation was performed alongside the RCT involving 3 intervention arms (see the section *Interventions: Content and Development Trajectory*). Following the definition given by the Medical Research Council, the aim of this process evaluation was to understand the functioning of a complex intervention consisting of multiple components [18]. The design of the RCT is described elsewhere [14].

Participants

Family caregivers as well as specialized dementia nurses participated in the process evaluation.

Inclusion criteria for family caregivers were the same as the criteria used in the RCT: family caregivers aged at least 18 years, who were a partner or relative of a person diagnosed with dementia who is living at home, having contact at least weekly with the person with dementia, with internet access and who provided online consent [14]. In total, 81 family caregivers participated in the RCT (major [n=27], medium [n=27], or minor [n=27] intervention arms).

Inclusion criteria for the specialized dementia nurses were: (1) a Bachelor's or Master's degree in nursing, and (2) advanced training in dementia care. In total, 4 nurses participated.

Interventions: Content and Development

Family caregivers were randomly allocated to 1 (major), 2 (medium), or 3 (minor) intervention arms.

1. The major intervention arm was the most comprehensive and consisted of the following 3 elements:
   - The first element consisted of email contacts with a specialized dementia nurse thrice. In the email contacts, the nurse helped family caregivers online to manage behavioral changes, guided by an intervention protocol developed by the project team members (JGH, ALF, PJV, and IvA), in consultation with the nurses themselves. The Dutch protocol (available on request from the first author) was structured using the steps of the 5A model of self-management support [19,20] and Kitwood's person-centered care theory [21]. The 5A model comprises the following 5 steps: assessing; advising; agreeing on goals; assisting in anticipating barriers and developing a specific action plan; and arranging follow-up [19,20,22,23].
   - Another element was providing online videos on how to manage the relative's behavior changes and how to improve self-efficacy in managing this behavior. There were 6 videos dealing with different common types of behavior changes: dependent behavior, suspicious behavior, aggressive behavior, apathy or indifference, restlessness at night, and masking behavior. Each video had the same structure, starting with possible causes and related solutions for responding or coping with the specific behavior, and ending by emphasizing that it is important that family caregivers take good care of themselves. Family caregivers could choose how many videos they watched depending on their personal needs and the behavioral changes encountered in their relative with dementia. The videos (along with the e-bulletins mentioned below—see element c) were developed by the coauthors BMW, IvA, and AMP, in close collaboration with colleagues from the Trimbos Institute and the Dutch Alzheimer's Society, family caregivers of people with dementia and other experts. In the first step of the development process, a desk search was performed to obtain an impression of what is known in the literature about methods of influencing behavior approached from a person-centered perspective [21] and how family caregivers experience different kinds of behavioral changes in their relative with dementia. Experts also provided input for various aspects of the videos (eg, cognitive behavioral therapy principles, persuasive communication, modeling, and active learning). Video scripts and pilot videos were tested by family caregivers at several points during development. The videos are available on https://dementie.nl/online-training.
   - Providing e-bulletins containing practical information about various types of behavioral changes and how to manage them was the third element. The same behavior changes were covered in the e-bulletins as in the videos. The e-bulletins included assignments that were designed to help caregivers interpret the generic information in the context of their own situation, to reflect on what might be causing the behavior changes, how they would like to cope with the behavior, and how they would like to respond. During the development process, the e-bulletins were tested together with the online videos. They have the same theoretical basis as the videos, and the people involved in the development of the videos were also involved in developing the e-bulletins.
2. The medium intervention, consisting only of the online videos and e-bulletins (elements b and c above);
3. The minor intervention, consisting only of the e-bulletins (element c).

Data Collection

A schematic overview of the data collection methods used is given in Table 1. In some parts of the process evaluation, the sample concerned all family caregivers participating in the RCT (n=81), whereas in other parts of the process evaluation, only subsamples participated.

As can be seen in the second column of Table 1, quantitative data involved nurses' records of the number of times personal email contacts occurred per family member, clicks on links to the online videos and e-bulletins, and evaluation questions answered by family caregivers in a questionnaire. The evaluation questions were part of the questionnaire used at the end of the RCT [15].

As shown in the third column of Table 1, the qualitative data were collected in semistructured interviews with family caregivers.

In the last questionnaire used at the end of the RCT, family caregivers were asked if they would like to take part in such an interview. In total, 41 family caregivers were willing to participate. Of these, 12 were purposively recruited with a spread of intervention arms and background characteristics (eg, gender, age, and living with or separately from the relative with dementia). They were sent an information letter by email and were asked to give their consent by email if they were willing to be interviewed. All interviews were conducted by telephone by one of the coauthors.

Semistructured interviews were also conducted with the 4 specialized dementia nurses who provided the personalized email contacts with the family caregivers. The topic list addressed how the nurses evaluated their support in the personal tailored email contacts. All interviews with the nurses were carried out by one researcher (IvA). Three interviews were conducted by telephone; one interview took place at the Trimbos Institute.

Finally, the content of email contacts was analyzed regarding family caregivers' request for help, referral to the online videos and nurses' use of the intervention protocol based on the 5A model.

Table 1. Data collection methods (quantitative and qualitative) used for each research question.

Research aims	Quantitative data	Qualitative data	Data collection period
To understand the actual use of family caregivers of the elements of the self-management support	Recording the actual use of personal email contact with nurse by 27 family caregivers	—[a]	March to August 2017
	Clicks on the video links by 54 family caregivers	—	March to August 2017
	Clicks on the e-bulletin links by 81 family caregivers	—	March to August 2017
To understand family caregivers' evaluation and satisfaction with the various elements of the online self-management support interventions	Evaluation questions in a survey with Likert scale, send to 81 family caregivers	—	March to August 2017
	—	Semistructured interviews with 12 family caregivers	July and August 2017
To understand nurses' usage and evaluations of the online support through tailored email contacts	—	Semistructured interviews with 4 nurses	September 2017
	—	Analysis of the content of email contacts between 27 family caregivers and nurse	March to August 2017

[a]Not applicable.

Data Analysis

Quantitative Data

Records and clicks on links were descriptively analyzed using Microsoft Excel (version 2010). The evaluation questions in the questionnaire were analyzed descriptively (frequencies and percentages) using SPSS software (IBM Corporation).

Qualitative Data

The transcribed verbatim audio-recorded interviews were analyzed independently by 2 researchers (JGH and IvA) using the principles of thematic analysis [24]. First, the researchers repeatedly read the data and looked for meanings and patterns in the data. Second, an initial list of codes was generated about what was in the data and what was interesting for the research questions. Third, the various codes were sorted into potential themes and then fourth, refined so that data within the themes fitted together meaningfully. Fifth, the themes were further refined by analyzing the data within the themes. Finally, once there was a set of fully detailed themes, the final analyses were written down [24]. Coding and interpretation of the codes were discussed at several moments in the analysis process by the researchers to reach consensus and to refine the analyses. In addition, interim and final analyses were also discussed with other coauthors. Furthermore, member checking was performed by discussing interim and final analyses with one of the nurses who was involved in the email contacts (PJV).

The content of email contacts between 27 family caregivers and the nurses was analyzed by 1 researcher (JGH) looking at their request for help, referral to the online videos and the use of the 5A model. A second reviewer (ALF) screened a random selection (email contacts of 3 family caregivers).

Ethical Considerations

The Medical Ethics Committee of the Vrije University Medical Center approved this study (reference 2016.559). All participating family caregivers and dementia nurses gave informed consent. All data were stored according to the rules of the Dutch Data Protection Act.

Results

The collected quantitative and qualitative data are categorized based on the 3 different elements of the intervention: email contacts with nurses, online videos, and e-bulletins.

Family Caregivers' Usage of Email Contacts With Nurses

A total of 27 family caregivers were assigned to the major intervention arm, meaning that they had the opportunity to have personal email contact with a dementia nurse in addition to the videos and e-bulletins. Of the 27 family caregivers, 21 (78%) actually made use of the opportunity. Almost half of the family caregivers (13/27, 48%) had email contacts thrice, 4 had twice (15%), and another 4 (15%) had once (Table 2).

Table 2. Data from the recording form for personal email contacts kept by the nurses.

Personal email contacts	Value
Family caregivers in the major intervention arm, n	27
Had email correspondence thrice	13
Had email correspondence twice	4
Had email correspondence once	4
No email correspondence	6
Total number of times the email contacts occurred	51
Time spent per email contact by nurses (min), mean (range)	35 (20-55)

Family Caregivers' Evaluation and Satisfaction With Email Contacts With Nurses

A total of 27 family caregivers assigned to the major intervention arm were asked to complete evaluation questions about the email contacts (Table 2, second column). A total of 16 family caregivers completed the evaluation questions and had email contact with a nurse. The majority (12/16; 75%) valued the personal email contacts with the nurses in addition to the videos and e-bulletins. The nurses' explanation and advice given in the email contacts were clear for most family caregivers (12/16, 75%), and more than half of them (9/16, 56%) said they could immediately use the nurses' advice in managing the behavior of their relative with dementia (Table 3).

A total of 4 family caregivers in the major intervention arm were interviewed about how they evaluated the personal email contacts with the nurse. They stated that they got the most out of these contacts, compared with online videos and e-bulletins, because of the personal aspect. They appreciated that they were given the opportunity to reflect with the nurse on how they were dealing with their relative's situation, which made them aware that they had to take a step back in some situations. They also liked the tips and ideas that the nurses gave them about how to act in their situation. In addition, they said that it was good to get confirmation that you were doing it correctly:

> *Now like I said: you talk. At least you are then communicating with somebody [the nurse] who understands what it's about. You don't have to keep on reinventing the wheel then, in fact. You can just say, well, I'm coming up against this and that. Oh—watch out for this, watch out for that. That's simply very pleasant, I reckon.* [Participant 10]

One family caregiver said that she did not use the email contacts. The reasons were not only the lack of time but also that the counselling by email was not attractive because you then have to put your emotions and questions on paper. That was a barrier for this family caregiver, who also said that the barrier would have been much lower if the counselling had been by phone.

Table 3. Evaluation questions on Likert scales.

Survey questions	Major (n[a]=16), n (%)	Medium (n=21), n (%)	Minor (n=15), n (%)
The personal email correspondence with the nurse added value to the video and e-bulletins			
Completely agree/agree	12 (75)	—[b]	—
Neutral	2 (12)	—	—
Disagree/completely disagree	2 (12)	—	—
The nurse's explanation and advice were clear			
Completely agree/agree	12 (75)	—	—
Neutral	3 (19)	—	—
Disagree/completely disagree	1 (6)	—	—
I was able to use the advice of the nurses immediately in managing the behavior of my relative with dementia			
Completely agree/agree	9 (56)	—	—
Neutral	6 (37)	—	—
Disagree/completely disagree	1 (6)	—	—
The videos and e-bulletins fitted my situation			
Completely agree/agree	10 (62)	9 (43)	—
Neutral	6 (37)	11 (52)	—
Disagree/completely disagree	0 (0)	1 (5)	—
The videos and e-bulletins helped me to manage the behavior of my relative with dementia			
Completely agree/agree	11 (69)	10 (48)	—
Neutral	5 (31)	10 (48)	—
Disagree/completely disagree	0 (0)	1 (5)	—
In addition to videos and e-bulletins, I would have liked to receive extra support from a nurse by email			
Completely agree/agree	—	6 (29)	—
Neutral	—	13 (62)	—
Disagree/completely disagree	—	2 (10)	—
The e-bulletins fitted my situation			
Completely agree/agree	—	—	7 (47)
Neutral	—	—	6 (40)
Disagree/completely disagree	—	—	2 (13)
The e-bulletins helped me to manage the behavior of my relative with dementia			
Completely agree/agree	—	—	8 (53)
Neutral	—	—	5 (33)
Disagree/completely disagree	—	—	2 (13)
In addition to the e-bulletins, I would have liked to receive extra support from a nurse by email			
Completely agree/agree	—	—	6 (40)
Neutral	—	—	6 (40)
Disagree/completely disagree	—	—	3 (20)

[a]Family caregivers who completed the evaluation questions.

[b]Data not applicable.

Family Caregivers' Usage of Online Videos

In total, 54 family caregivers (27 in the major intervention arm and 27 in the medium intervention arm) had access to 6 videos about how to manage behavioral changes in their relative with dementia. Of them, 43 (80%) clicked at least 1 video. Clicks on the videos are listed in Table 4.

Table 4. Clicks on the links to the videos and e-bulletins.

Clicks on videos and e-bulletins	n (%)
Total number of family caregivers who clicked videos	43 (80)
Family caregivers in the major intervention arm who watched at least one video (n=27)	22 (81)
Family caregivers in the medium intervention arm who watched at least one video (n=27)	21 (78)
Total clicks on e-bulletins	30 (37)
Family caregivers in the major intervention arm who watched at least one e-bulletin (n=27)	5 (19)
Family caregivers in the medium intervention arm who watched at least one e-bulletin (n=27)	7 (26)
Family caregivers in the minor intervention arm who watched at least one e-bulletin (n=27)	18 (67)

Family Caregivers' Evaluation and Satisfaction With the Videos and E-Bulletins

A total of 54 family caregivers were asked to complete evaluation questions about the videos (Table 3, second column). A total of 37 caregivers watched at least 1 video (16 in the major intervention arm and 21 in the medium intervention arm). Half of them (both major and medium arms, 19/37, 51%) said that the videos and e-bulletins fitted their personal situation, and more than half stated that they helped them to better manage with the behavior of the person with dementia (21/37, 57%).

In total, 9 family caregivers who had access to the videos (4 in the major intervention arm and 5 in the medium intervention arm) were interviewed about how they evaluated the videos. They said that they thought the videos were well-structured and pleasant to watch. They also found the content clear and useful. The tips given in them were reckoned to be useful; watching the videos gave them new ideas for coping with the behavioral changes in their relative:

> *Well, because it's important for you to have a clear picture as well. It's useful to know what I ought to be doing. That really does help quite a bit. Otherwise there's a lot of conflict and so forth, instance defiance or whatever—quite a lot. It lets you know how to tackle the situation: let's put it like that.* [Participant 5]

Some of the family caregivers found the content and the stories of other family caregivers recognizable and helpful. Others said they did not relate to much that was in the videos because there was no change in behavior in their situation or the behavior was expressed differently. They said that this meant the videos were less useful to them:

> *With my husband, it was mostly about the aggression and waking up at night and that wasn't something I really saw in the video or in the text. And that was what I find so typical. There were a few bits in that I recognized, but I didn't get the feeling that the situation really fitted in very well with ours.* [Participant 4]

One family caregiver also remarked that the videos and the e-bulletins were suitable primarily in the early phases of dementia; another said that the information was too sketchy for family-based caregivers dealing with dementia in its later stages.

Family Caregivers' Usage of E-Bulletins

All family caregivers (n=81) had access to the e-bulletins (through a link). The e-bulletins contained practical information about various types of behavioral changes and tasks to help reflect on their possible causes and how to influence and cope with them. In total, 30 family caregivers out of 81 (37%) clicked at least 1 e-bulletin. In the minor intervention arm, the percentage who clicked the e-bulletins was the highest (18/27, 67%; Table 3).

Family Caregivers' Evaluation and Satisfaction With E-Bulletins

Of 27 family caregivers in the minor intervention arm, 15 (56%) answered the evaluation questions (Table 3, fourth column). Almost half (7/15, 46%) said that the e-bulletins fitted their situations and that the e-bulletins helped them to manage behavioral changes in the person with dementia (8/15, 53%).

In total, 12 family caregivers (4 in the major intervention arm, 5 in the medium intervention arm and 3 in the minor intervention arm) were interviewed on how they evaluated the e-bulletins. A number of family caregivers stated that the information in the e-bulletins was clear and recognizable as well as being helpful to read again. Some also said that one of the benefits was that there was one e-bulletin of each type of behavior. Conversely, others felt that the content of the e-bulletins was not always recognizable and that they were unable to translate it well to their own situations. One family caregiver said that the e-bulletins were not concrete enough, and she also perceived the e-bulletins as a bit patronizing at times. On the other hand, this family caregiver also said that this might be just a personal opinion.

Some of the group who had also seen the video felt that the e-bulletin was a good addition to the videos, whereas others set more store by it because information from the videos was enough for them.

The family caregivers who only received the e-bulletins mostly thought they were informative, although one family caregiver said that information did not help in her situation. Others said that the information meant they were more aware of what they could come up against and that it put a different perspective on the behavior for them. Moreover, understanding the behavior better because of the information from the e-bulletins let them be more patient in dealing with the behavior:

> *Explaining the behavior and how you have to respond to it, right? Most of the time you have to count to ten first or—as I always say—sometimes a hundred. Like that.* [Participant 3]

The family caregivers would recommend the e-bulletins to others. One of them advised distributing this information among professionals too, having noted that they do not always know enough about behavioral changes.

Nurses' Evaluation and Satisfaction With Providing Tailored Email Contacts

Four specialized dementia nurses provided online self-management support via email. In total, the nurses had email contacts with family caregivers 51 times. The time spent by the nurses varied from 20 min to 55 min (mean 35 min) per email contact (Table 1).

Semistructured interviews were held with 4 nurses to get an idea of their use and evaluations of providing online self-management support. Categorization resulted in 4 themes: background characteristics and expectations of family caregivers, evaluation of the online assistance, evaluation of the intervention protocol with the 5A model, number of times the email contacts, and the perceived effect.

Background Characteristics and Expectations of Family Caregivers

Two specialist nurses said that the family caregivers had partners in an advanced stage of dementia. One nurse said that she got the impression that the family caregivers were overloaded. Moreover, the nurses noted that some of the caregivers had one or more people helping them and were deliberately busy collecting information about the condition. In addition, a nurse said that the family caregivers were not aware that they were also tackling their situations too.

In terms of the expectations of the family caregivers, the majority of the nurses had the impression that family caregivers were looking for a release valve and a listening ear. A number of the caregivers needed concrete ideas about how to deal with behavioral changes in their relative. One nurse also said that she noticed that she was being asked questions about case management, for instance about coordinating care for the relative.

Evaluation of the Online Assistance

The nurses said that there were pros and cons to giving online assistance. One nurse said that putting the situation down on paper was one of the benefits of online counselling because the family caregivers then got a better picture of the severity, and the situation would sink in more quickly:

> *Yes [...] because the family caregivers are e-mailing and putting things into words, the seriousness of the problem is made a bit clearer, I reckon. I get that idea quite strongly. Putting it on paper can in fact point out the severity—almost as if they're saying they can't cope any more. Yes, that does help. It paints a picture of the changed behavior, and shows that action is needed as well.* [Nurse 1]

They also felt it was an advantage that you can ask encouraging questions, but the nurse wondered whether this matched the family caregivers' expectations of this online assistance.

Giving online counselling was also felt to be *awkward* because you cannot look anyone in the eye, and it is then more difficult to assess the situation. They found it tricky to get the right tone for approaching the family caregiver. As the counselling was online, the nurse did not know if the advice had been understood by the family caregiver. If the caregiver no longer responded, the nurse did not know if they had said something wrong or if there was another, unrelated reason.

Another nurse said that online assistance is suitable for practical questions, but that you need more time and need to know more in the role of health care provider if it is about people being overburdened or about changed behavior. Another nurse believed that it became easier as you did it more often. A certain amount of practice is needed if this counselling is to be provided properly.

Evaluation of the Intervention Protocol With the 5A Model

One nurse said that the 5A model could help a lot in the online counselling but that the nurses had difficulties with the application of the model. The link between the video content and the 5A model was also unclear. The reason was that they had a feeling that the family caregivers needed other assistance, for example, providing a listening ear. The videos and email counselling focused on coping with the changed behavior, but the nurses noticed that the family caregivers had more of a need to talk about things. Getting them to talk about the behavioral changes and think about them felt like the nurses were pushing.

Number of Times Email Contacts Occurred

Opinions varied as to whether the number of times the email contacts occurred was sufficient. Two of the nurses said that it was enough. One nurse did state as a condition that the contacts should then only focus on the behavioral changes and not on other questions and advice. Another nurse doubted whether contacts occurring 3 times was enough to have an effect. The emails from the family caregivers contained a lot of information, not only about to change behavior but also about the other problems involved. Another nurse said she got the impression that family caregivers enjoyed watching the videos but did not think that they actually wanted to do anything as result.

The Perceived Effect

Most of the nurses said that their assistance meant that family caregivers could get things off their chest or that the family caregivers felt they had been listened to. One nurse said that it was a help that the family caregivers had taken a moment to think about the behavioral changes in their relative with dementia. She also thought that the tips she had given about how to make thorny subjects open to discussion had helped. According to one nurse, effective elements were the attention paid to the personal situations of the family caregivers and being able to reflect on them together. This nurse was also able to give the family caregivers tips about other forms of assistance. Another nurse believed that the email contacts had helped the

family caregivers translate what was happening in the videos to their own situations. In the case of one family caregiver, a nurse had the impression that the counselling had no effect because the person in question was already so overburdened that email contact was too much. Another nurse did not believe that it had given the family caregivers a better picture of behavioral changes because the nurses did not have the right skills for online counselling and because the need for assistance among family caregivers was so diverse.

Analysis of the Content of Email Contacts Between Family Caregiver and Nurse

Of 27 family caregivers (78%) 21 had email contact with a nurse. As data were missing for 2 family caregivers, email contacts from 19 family caregivers were analyzed. Of the 19 family caregivers, 11 (58%) did not express an explicit goal in the email contacts. In 15 cases (79%), the content of the emails was about behavioral changes in their relative with dementia. A total of 4 family caregivers (21%) also discussed about caregiving stress. A total of 6 (32%) family caregivers discussed other caregiving issues not related to behavioral changes. In 5 cases (26%), the nurse referred to the online videos in the email contacts (Table 5).

The first step in the 5A model (*assessing*) was used by the nurses in all email contacts. The second step (*advising*) was used in about half of the cases. The other steps of the 5A model (*agreeing on goals*, *assisting in anticipating barriers and developing a specific action plan*, and *arranging follow-up*) hardly occurred at all in the email contacts.

Table 5. Content of all email contacts between family caregivers and a nurse (n=19).

Email content	n (%)
Explicitly formulated request for help	
Yes	8 (42)
No	11 (58)
Content discussed in 1 or more emails	
Behavioral changes	15 (79)
Managing caregiver stress	4 (21)
Other caregiving issues (other than behavioral changes of the relative with dementia)	6 (32)
A link to the online videos	
Yes	5 (26)
No	14 (74)

Discussion

Principal Findings

Through this process evaluation, we aimed to gain an idea of (1) actual use by family caregivers of the various elements of online self-management support, (2) family caregivers' evaluation and satisfaction with the various elements, and (3) nurses' usage and evaluations of the online support through the tailored email contacts. This process evaluation was performed alongside an RCT [14] in which the effectiveness was studied of an online self-management support intervention involving tailored email contacts with a specialized dementia nurse combined with online videos and e-bulletins. Contrary to our expectations, no statistical evidence was found for the major and medium online self-management interventions compared with minor intervention (involving e-bulletins only) on family caregivers' self-efficacy [15]. Although no effects were found, this evaluation noted that family caregivers valued the email contacts with the specialist nurse. They mentioned that receiving confirmation from a professional that they were doing the right thing was really important to them. Previous studies also found that being acknowledged by professionals and peers for the everyday care they provided is extremely important for family caregivers in helping them cope with daily challenges [8,25]. They also felt that the email contacts offered added value above the videos and e-bulletins. Family caregivers who received the videos and e-bulletins mentioned difficulties in translating the information and advice to their own situations. It could therefore be suggested that an online personal approach is needed to acknowledge the highly complex situation of family caregivers and subsequently assist them by providing tailored online information and advice for their personal situations.

This process evaluation also suggests possible explanations for the unexpected results in the RCT by understanding how the intervention was used and was evaluated by the people involved. First, this process evaluation showed variation in the extent that family caregivers made use of the various elements of the online self-management support. 78% of family caregivers had an email contact with a nurse (21/27), 80% watched 1 or more online videos (43/54), and 37% clicked an e-bulletin (30/81). The distinction between the 3 intervention arms consequently becomes less, making it difficult to demonstrate effects [17]. Nonuse of an intervention is a methodological known difficulty in online trials and may explain why interventions fail to show a measurable effect for the intervention [17,26,27]. For electronic health (eHealth) interventions to present an effect, they need to be accepted and used in the intended way to benefit the participants the most [28]. However, improving the use of eHealth interventions is complex, and more insights are needed for investigating the reasons why people accept, adopt, and adhere to eHealth interventions so that their behavior can be influenced [28].

Second, according to the nurses, the participants involved in the email contacts were mainly family caregivers who availed the services of 1 or more health professionals and were highly engaged in collecting information about dementia. This group would then already have information and advice on how to cope with behavioral changes, which might explain why family caregivers wanted to share their stories and express their feelings instead of finding other ways to self-manage the behavioral changes of their relative with dementia. For future research, it is important to determine which family caregivers will benefit most from what type of support. This would provide insight that can be used to provide the intervention in a more cost-effective way. This, for example, means that nurses' support can be provided to the people who are likely to benefit most.

Another possible explanation for finding no statistical evidence for the benefits of email contact between family caregivers and nurses (combined with videos and e-bulletins) may be how the intervention was carried out. In many cases, only the first two A's (*assessing* and *advising*) were completed. Using the 5A model turned out to be difficult as it was new to the nurses. Previous studies' results were comparable, as the last 2 A's (*assist* and *arrange*) seem to be delivered least often by nurses [19,29,30]. However, those components are most important for producing meaningful and lasting behavioral changes [19]. Future research therefore needs to investigate how all steps of the 5A model could be performed online.

When providing online support, the dosage of online intervention should also be considered. It is for instance striking that only a few (37%, 30/81) family caregivers clicked e-bulletins. This could be explained by the fact that not everything that is offered will also be used. This may be illustrated by the low usages rates of the e-bulletins by family caregivers who also had email contact with a nurse and access to the online videos. This indicates that informal caregivers do not stick to the intervention but decide for themselves what care is needed and fits their unique situation. Tailored information and advice should therefore be offered in a way that is geared to family caregivers' needs [31]. This could include a differentiated offer of support instead of offering multiple kinds of support. This enables family caregivers get help that is based on their needs, self-management abilities, and home situations.

Strengths and Limitations

The mixed methods design combining quantitative and qualitative data enabled better understanding of how online self-management support interventions were used and evaluated by both the family caregivers and dementia nurses involved. Using telephone interviews let family caregivers participate without extra traveling time for the family caregiver or researcher. The information gathered can be used to develop online self-management support further for families facing dementia. Furthermore, the validity of the results was enhanced by combining quantitative and qualitative data [32]. However, findings of this study need to be considered within the context of a number of methodological limitations. First, tracked usage data were measured in clicks that represent page views. People who click a link do not however necessarily watch the whole online video or read the e-bulletin. The numbers found could therefore overestimate family caregivers' utilization of an intervention component. Click data should therefore be seen in combination with other evaluation methods [33]. Second, no data were collected for the 6 family caregivers who did not use the email contacts. Barriers preventing family caregivers from making email contact with a nurse could therefore potentially have not been detected.

Conclusions

There was a variation in the extent to which family caregivers utilized the various elements of online self-management support. They valued the tailor-made approach in the email contacts. According to the nurses involved, online personal email contacts was mostly used to express feelings concerning coping with changing behavior. Nurses' usage and evaluations of providing self-management support online were mixed, as it is a relatively new task for nurses. More research is needed to investigate the reasons why people accept, adopt, and adhere to online interventions to reduce nonuse and to support them appropriately by providing tailored (online) information and advice for their personal situations.

Acknowledgments

This study is part of the research program *NurseSMS, Nurses Self-management Support for people facing progressive incurable diseases* (2014-2019), which is funded by ZonMw; the Netherlands Organization for Health Research and Development, Program *Tussen Weten en Doen*. The online videos and e-bulletins were developed with financial support from the NuntsOhra fund, the Sluyterman van Loo fund, Stichting RCOAK, and Dutch Alzheimer's Society.

Conflicts of Interest

None declared.

References

1. Egilstrod B, Ravn MB, Petersen KS. Living with a partner with dementia: a systematic review and thematic synthesis of spouses' lived experiences of changes in their everyday lives. Aging Ment Health 2019 May;23(5):541-550. [doi: 10.1080/13607863.2018.1433634] [Medline: 29405735]
2. Chiao CY, Wu HS, Hsiao CY. Caregiver burden for informal caregivers of patients with dementia: a systematic review. Int Nurs Rev 2015 Sep;62(3):340-350. [doi: 10.1111/inr.12194] [Medline: 26058542]

3. Feast A, Orrell M, Charlesworth G, Melunsky N, Poland F, Moniz-Cook E. Behavioural and psychological symptoms in dementia and the challenges for family carers: systematic review. Br J Psychiatry 2016 May;208(5):429-434 [FREE Full text] [doi: 10.1192/bjp.bp.114.153684] [Medline: 26989095]
4. Müller-Spahn F. Behavioral disturbances in dementia. Dialogues Clin Neurosci 2003 Mar;5(1):49-59 [FREE Full text] [Medline: 22034255]
5. Huis in het Veld J, Verkaik R, van Meijel B, Verkade PJ, Werkman W, Hertogh C, et al. Self-management by family caregivers to manage changes in the behavior and mood of their relative with dementia: an online focus group study. BMC Geriatr 2016 May 3;16:95 [FREE Full text] [doi: 10.1186/s12877-016-0268-4] [Medline: 27142664]
6. Zwaanswijk M, Peeters JM, van Beek AP, Meerveld JH, Francke AL. Informal caregivers of people with dementia: problems, needs and support in the initial stage and in subsequent stages of dementia: a questionnaire survey. Open Nurs J 2013;7:6-13 [FREE Full text] [doi: 10.2174/1874434601307010006] [Medline: 23346266]
7. Barlow J, Wright C, Sheasby J, Turner A, Hainsworth J. Self-management approaches for people with chronic conditions: a review. Patient Educ Couns 2002;48(2):177-187. [doi: 10.1016/S0738-3991(02)00032-0] [Medline: 12401421]
8. Huis in het Veld J, Verkaik R, van Meijel B, Verkade PJ, Werkman W, Hertogh C, et al. Self-management support and ehealth when managing changes in behavior and mood of a relative with dementia: an asynchronous online focus group study of family caregivers' needs. Res Gerontol Nurs 2018 May 1;11(3):151-159. [doi: 10.3928/19404921-20180216-01] [Medline: 29498746]
9. Alleyne G, Hancock C, Hughes P. Chronic and non-communicable diseases: a critical challenge for nurses globally. Int Nurs Rev 2011 Sep;58(3):328-331. [doi: 10.1111/j.1466-7657.2011.00912.x] [Medline: 21848778]
10. Hopwood J, Walker N, McDonagh L, Rait G, Walters K, Iliffe S, et al. Internet-based interventions aimed at supporting family caregivers of people with dementia: systematic review. J Med Internet Res 2018 Jun 12;20(6):e216 [FREE Full text] [doi: 10.2196/jmir.9548] [Medline: 29895512]
11. Pot AM, Blom MM, Willemse BM. Acceptability of a guided self-help internet intervention for family caregivers: mastery over dementia. Int Psychogeriatr 2015 Aug;27(8):1343-1354. [doi: 10.1017/S1041610215000034] [Medline: 25648589]
12. Parra-Vidales E, Soto-Pérez F, Perea-Bartolomé MV, Franco-Martín MA, Muñoz-Sánchez JL. Online interventions for caregivers of people with dementia: a systematic review. Actas Esp Psiquiatr 2017 May;45(3):116-126 [FREE Full text] [Medline: 28594057]
13. Boots LM, de Vugt ME, van Knippenberg RJ, Kempen GI, Verhey FR. A systematic review of internet-based supportive interventions for caregivers of patients with dementia. Int J Geriatr Psychiatry 2014 Apr;29(4):331-344. [doi: 10.1002/gps.4016] [Medline: 23963684]
14. Huis in het Veld J, Willemse B, van Asch IF, Zwaaftink RB, Verkade PJ, Veldhuijzen NJ, et al. Online self-management support for family caregivers to help them manage behavior changes in their relative with dementia: study protocol for a randomized controlled trial and a process evaluation. JMIR Res Protoc 2017 Nov 28;6(11):e239 [FREE Full text] [doi: 10.2196/resprot.8365] [Medline: 29183870]
15. Huis in het Veld J, Willemse BM, van Asch IFM, Groot Zwaaftink RBM, Verkade PJ, Twisk JWR, et al. Effects of online self-management support of family caregivers to deal with behavior changes of the relative with dementia (part 2): a randomized controlled trial. J Med Internet Res (forthcoming). [doi: 10.2196/13001]
16. Eysenbach G. Consort-eHealth: implementation of a checklist for authors and editors to improve reporting of web-based and mobile randomized controlled trials. Stud Health Technol Inform 2013;192:657-661. [doi: 10.3233/978-1-61499-289-9-657] [Medline: 23920638]
17. Eysenbach G. The law of attrition. J Med Internet Res 2005 Mar 31;7(1):e11 [FREE Full text] [doi: 10.2196/jmir.7.1.e11] [Medline: 15829473]
18. Moore G, Audrey S, Barker M, Bond L, Bonell C, Hardeman W, et al. Process evaluation of complex interventions: Medical Research Council guidance. Br Med J 2015 Mar 19;350:h1258 [FREE Full text] [doi: 10.1136/bmj.h1258] [Medline: 25791983]
19. Glasgow RE, Emont S, Miller DC. Assessing delivery of the five 'As' for patient-centered counseling. Health Promot Int 2006 Sep;21(3):245-255. [doi: 10.1093/heapro/dal017] [Medline: 16751630]
20. Whitlock EP, Orleans CT, Pender N, Allan J. Evaluating primary care behavioral counseling interventions: an evidence-based approach. Am J Prev Med 2002 May;22(4):267-284. [Medline: 11988383]
21. Kitwood T. Dementia Reconsidered: The Person Comes First. University of Michigan, USA: Open University Press; 1997.
22. Wagner EH, Austin BT, Davis C, Hindmarsh M, Schaefer J, Bonomi A. Improving chronic illness care: translating evidence into action. Health Aff (Millwood) 2001;20(6):64-78. [doi: 10.1377/hlthaff.20.6.64] [Medline: 11816692]
23. Glasgow RE, Davis CL, Funnell MM, Beck A. Implementing practical interventions to support chronic illness self-management. Jt Comm J Qual Saf 2003 Nov;29(11):563-574. [doi: 10.1016/S1549-3741(03)29067-5] [Medline: 14619349]
24. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]

25. van Wijngaarden E, van der Wedden H, Henning Z, Komen R, The AM. Entangled in uncertainty: the experience of living with dementia from the perspective of family caregivers. PLoS One 2018;13(6):e0198034 [FREE Full text] [doi: 10.1371/journal.pone.0198034] [Medline: 29897922]
26. Eysenbach G. Issues in evaluating health websites in an internet-based randomized controlled trial. J Med Internet Res 2002 Dec;4(3):E17 [FREE Full text] [doi: 10.2196/jmir.4.3.e17] [Medline: 12554548]
27. Kelders SM, Kok RN, Ossebaard HC, van Gemert-Pijnen JE. Persuasive system design does matter: a systematic review of adherence to web-based interventions. J Med Internet Res 2012 Nov 14;14(6):e152 [FREE Full text] [doi: 10.2196/jmir.2104] [Medline: 23151820]
28. van Gemert-Pijnen JE, Kelders SM, Kip H, Sanderman R, editors. eHealth Research, Theory and Development: A Multi-Disciplinary Approach. Abingdon: Routledge; 2019.
29. Glasgow RE, Eakin EG, Fisher EB, Bacak SJ, Brownson RC. Physician advice and support for physical activity: results from a national survey. Am J Prev Med 2001 Oct;21(3):189-196. [doi: 10.1111/1753-6405.12455] [Medline: 11567839]
30. Slev VN, Pasman HR, Eeltink CM, van Uden-Kraan CF, Verdonck-de Leeuw IM, Francke AL. Self-management support and ehealth for patients and informal caregivers confronted with advanced cancer: an online focus group study among nurses. BMC Palliat Care 2017 Nov 21;16(1):55 [FREE Full text] [doi: 10.1186/s12904-017-0238-4] [Medline: 29162081]
31. Huis in het Veld J, Verkaik R, Mistiaen P, van Meijel B, Francke AL. The effectiveness of interventions in supporting self-management of informal caregivers of people with dementia: a systematic meta review. BMC Geriatr 2015 Nov 11;15:147 [FREE Full text] [doi: 10.1186/s12877-015-0145-6] [Medline: 26561236]
32. Polit DF, Beck CT. Nursing Research: Generating and Assessing Evidence for Nursing Practice. Ninth Edition. Philadelphia: Lippincott Williams and Wilkins; 2012.
33. Song MJ, Ward J, Choi F, Nikoo M, Frank A, Shams F, et al. A process evaluation of a web-based mental health portal (WalkAlong) using Google analytics. JMIR Ment Health 2018 Aug 20;5(3):e50 [FREE Full text] [doi: 10.2196/mental.8594] [Medline: 30126832]

Abbreviations

eHealth: electronic health
RCT: randomized controlled trial

Edited by G Eysenbach; submitted 30.11.18; peer-reviewed by B Davis, N Davies, L de Souza; comments to author 21.03.19; revised version received 07.06.19; accepted 05.07.19; published 11.10.19.

Please cite as:
Huis in het Veld JG, van Asch IFM, Willemse BM, Verkade PJ, Pot AM, Blom MM, Groot Zwaaftink RBM, Francke AL
Process Evaluation of Nurse-Led Online Self-Management Support for Family Caregivers to Deal With Behavior Changes of a Relative With Dementia (Part 1): Mixed Methods Study
J Med Internet Res 2019;21(10):e13002
URL: https://www.jmir.org/2019/10/e13002
doi:10.2196/13002
PMID:31605517

Original Paper

Development and Efficacy of an Electronic, Culturally Adapted Lifestyle Counseling Tool for Improving Diabetes-Related Dietary Knowledge: Randomized Controlled Trial Among Ethnic Minority Adults With Type 2 Diabetes Mellitus

Kathleen Abu-Saad[1], PhD; Havi Murad[2], PhD; Rivka Barid[2,3], MA; Liraz Olmer[2], MSc; Arnona Ziv[4], MBA; Nuha Younis-Zeidan[5], RD, PhD; Vered Kaufman-Shriqui[6], RD, PhD; Michal Gillon-Keren[7], RD, PhD; Shmuel Rigler[8], MD; Yakir Berchenko[2,9], PhD; Ofra Kalter-Leibovici[1,10], MD

[1]Cardiovascular Epidemiology Unit, Gertner Institute for Epidemiology and Health Policy Research, Ramat Gan, Israel
[2]Biostatistics and Biomathematics Unit, Gertner Institute for Epidemiology and Health Policy Research, Ramat Gan, Israel
[3]Israel Central Bureau of Statistics, Jerusalem, Israel
[4]Information and Computerization Unit, Gertner Institute for Epidemiology and Health Policy Research, Ramat Gan, Israel
[5]Diet and Nutrition Service Unit for the Arab population in Sharon-Shomron District, Clalit Health Services, Arara, Israel
[6]Department of Nutritional Sciences, Ariel University, Ariel, Israel
[7]Institute of Endocrinology and Diabetes, Schneider Children's Medical Center, Petah Tikva, Israel
[8]Sharon-Shomron District, Clalit Health Services, Hadera, Israel
[9]Department of Industrial Engineering and Management, Ben-Gurion University of the Negev, Beer Sheva, Israel
[10]Sackler Faculty of Medicine, Tel-Aviv University, Tel Aviv, Israel

Corresponding Author:
Kathleen Abu-Saad, PhD
Cardiovascular Epidemiology Unit
Gertner Institute for Epidemiology and Health Policy Research
Sheba Medical Center
Ramat Gan, 52621
Israel
Phone: 972 526896782
Fax: 972 35349607
Email: kathleena@gertner.health.gov.il

Abstract

Background: Ethnic minority populations exhibit disproportionately high rates of type 2 diabetes mellitus (T2DM). Electronic health tools have the potential to facilitate the cultural adaptation and tailoring of T2DM education to improve the knowledge and management of diabetes mellitus (DM).

Objective: This study aimed (1) to develop an adaptable Interactive Lifestyle Assessment, Counseling, and Education (I-ACE) software to support dietitian-delivered lifestyle counseling among low-socioeconomic status (SES) ethnic minority patients with T2DM and (2) to evaluate its effect on DM-related dietary knowledge and management compared with standard lifestyle advice (SLA) in a randomized controlled trial (RCT).

Methods: The I-ACE software, developed in consultation with clinical dieticians, incorporates evidence-based dietary and physical activity (PA) recommendations and educational materials. The features and behavioral change techniques include quantitative lifestyle (dietary intake and PA) assessment and simulation, individually tailored education and recommendations, motivational interviewing, and goal setting. For the unblinded pilot RCT, 50 overweight or obese Arab adults (aged 40-62 years) with poorly controlled T2DM were recruited from primary care clinics and randomly assigned to receive 4 in-person, dietician-delivered counseling sessions over 6 months using either (1) the I-ACE tool (experimental arm) or (2) the SLA methods (comparison arm). All outcome assessments were face-to-face. DM-related dietary knowledge (primary outcome) was measured at baseline, 3, 6, and 12 months. Lifestyle and other parameters were measured before, during, and after the intervention. Multiple

linear regression and repeated measures linear mixed models were used to compare the changes in study outcomes and explore time trends in between-group and within-group changes.

Results: A total of 25 participants were enrolled in each arm, of whom 24 and 21 completed the final assessment of the primary outcome in the I-ACE and SLA arms, respectively. DM-related lifestyle knowledge increased more rapidly in the I-ACE arm than in the SLA arm (*P* value for study arm×time interaction=.02). Within the I-ACE arm, the mean (SE) differences in added sugar and dietary fiber intakes from baseline to 12 months were −2.6% (SE 1.0%) of total energy (*P*=.03) and 2.7 (SE 0.0) g/1000 kcal (*P*=.003), respectively. The odds of engaging in any leisure PA at 12 months tended to be higher in the I-ACE arm versus SLA arm, but did not reach statistical significance (odds ratio 2.8; 95% CI 0.7-11.6; *P*=.16). Both arms exhibited significant reductions in HbA_{1c} (*P* value for change over time <.001).

Conclusions: The use of the I-ACE software in a 6-month, 4-session dietician-delivered lifestyle counseling intervention improved the efficiency of lifestyle education, compared with SLA, among low-SES, ethnic minority patients with T2DM. This pilot trial provides justification for conducting a large-scale trial to evaluate its effectiveness and applicability in routine clinical care among ethnically diverse populations.

Trial Registration: ClinicalTrials.gov NCT01858506; https://clinicaltrials.gov/ct2/show/NCT01858506.

(J Med Internet Res 2019;21(10):e13674) doi:10.2196/13674

KEYWORDS

diabetes mellitus, type 2; diabetes-related dietary knowledge; lifestyle; software; culturally congruent care; ethnic minorities

Introduction

Background

Diabetes mellitus (DM) is a progressive chronic disease that can result in serious short- and long-term complications. Patient self-management education and support are fundamental to improving DM management [1-6], and guidelines recommend that every person with DM receive self-management education in a format appropriate for the patient's specific cultural, socioeconomic, literacy, and numeracy characteristics [2,3]. A growing number of self-management education programs for patients with DM are incorporating information technologies (IT) to improve their effectiveness and reach [7-9].

Nutrition therapy is one of the most challenging components of self-management for many patients with type 2 DM (T2DM). It is therefore particularly important that patients receive dietary education and collaborate with providers to develop individualized eating plans that are both implementable and sustainable and incorporate their preferences and needs [3]. Despite its critical role, few studies have focused on the nutritional education and counseling component of these DM self-management interventions [4-6,9-12] or reported their effectiveness in improving DM-specific dietary knowledge [7,8].

Patient-centered DM dietary education is especially critical for ethnic/racial minority populations, who often bear a disproportionately high burden of T2DM [2,13,14]. The standard dietary education and advice provided by mainstream health care services may not adequately address the daily challenges faced by minority patients as their cultural, social, and dietary norms and socioeconomic realities differ from that of the majority population. IT tools can provide new opportunities to make DM-related dietary education and counseling more relevant and individually tailored for patients [7,8]; however, few existing examples/initiatives have included sizeable proportions of ethnic minority patients [7,15]. There is a need to expand the evidence base for new digital health technologies that can address these needs [16] among the highest-risk, most vulnerable patient populations.

Objectives

In this paper, we report the development of a dietician-operated, culturally adaptable Interactive Lifestyle Assessment, Counseling, and Education (I-ACE) software. We further report the results of a pilot randomized controlled trial (RCT) evaluating its effect, compared with standard lifestyle advice (SLA), on improving DM-related dietary knowledge, lifestyle behaviors, and glycemic control in a sample of low-socioeconomic status (SES) adults with T2DM from the Arab minority in Israel.

Methods

Design

This was an open, parallel-group, pilot RCT randomized controlled pilot trial testing the effect of a 6-month, dietician-delivered, face-to-face diabetes lifestyle (diet and physical activity [PA]) counseling program using the I-ACE software compared with SLA. The I-ACE software was adapted to provide culturally congruent lifestyle counseling to Arab adults with T2DM.

Ethical Considerations

Ethics approval was obtained from the Helsinki committees of Sheba Medical Center and Clalit Health Services, and all participants provided written informed consent before enrollment. The study was registered at ClinicalTrials.gov (NCT01858506). The CONSORT eHealth checklist is provided in Multimedia Appendix 1.

Study Population and Participants

The Arab minority in Israel is an indigenous population that accounts for 20% of the total population. It differs in language, culture, and religion from the majority Jewish population and, for the most part, resides in residentially segregated and economically deprived communities [17]. It is characterized by

a lower SES and higher rates of T2DM, poor glycemic control, and diabetic complications, than the majority population [17-20].

A total of 50 eligible participants were recruited from the clinics of Clalit Health Services in 2 Arab towns. Inclusion criteria were: (1) age between 40 and 64 years, (2) diagnosis of T2DM, (3) having T2DM for ≤10 years, (4) body mass index (BMI) of 27 to 43 kg/m^2, and (5) hemoglobin A_{1c} (HbA_{1c}) between 8.0% and 11.3%. Participants were not eligible if they (1) were receiving short-acting insulin treatment, (2) had inadequate control of comorbid conditions, or (3) had factors that would limit adherence to interventions (eg, any medical or physical condition that prohibited participation in PA or following standard diets recommended for people with diabetes, pregnancy, uncontrolled psychiatric condition, significant cognitive impairment, or blindness).

Recruitment was conducted at the local Clalit Health Service clinics, in collaboration with the physicians of the potential participants. The lists of potential participants meeting the inclusion criteria were generated from Clalit electronic medical records and reviewed by the patients' physicians to identify eligible candidates. Patients with an HbA_{1c} result measured more than 3 months before eligibility screening were requested to get the test repeated. Baseline evaluation (including initial lifestyle knowledge, dietary and PA assessments, and anthropometric measurements) were completed in the clinic before randomization.

Eligible participants were randomly assigned in 1:1 ratio either to the I-ACE or SLA study arm using a permuted block design, central computer–generated randomization process, with even-numbered blocks of 2 to 6 participants. The randomization was performed at the Gertner Institute. Allocation concealment was maintained until after the provision of informed consent and randomization. Although the intervention type was known to the participants, dieticians, and study coordinator, the study statisticians were blinded to group allocation until the primary study outcomes analyses were completed. Group assignment was masked from all health care service providers other than the dietician.

The recruitment ran from August 2014 to January 2015. The participant follow-up and data collection from the electronic medical records was completed in March 2016. The pilot trial ended, according to the protocol, after all the participants had been followed up for 1 year.

Sample Size Considerations

The sample size calculation was based upon the reported differences in the nutritional knowledge change (percentage of correct responses) between the intervention and comparison groups in 2 nutritional educational interventions, which ranged from 8.9% to 11.5% [21,22]. A sample of 50 participants was needed to provide 90% power to detect a statistically significant difference of this magnitude at the 5% level using a 2-sided test between the experimental and comparison groups, allowing for a dropout rate of 10%.

Interventions

The study included 2 active intervention arms: (1) the I-ACE experimental arm and (2) the SLA comparison arm. The experimental arm of the intervention used the I-ACE software.

Information Technology Tool Description

The I-ACE software was designed for use by dieticians to support and enrich a patient-centered clinical lifestyle counseling process. It is a multifeatured tool that supports collecting data on habitual dietary and PA behaviors and using these data to calculate actionable, graphically displayed summary measures (eg, average daily or weekly food/nutrient intakes and PA). Additional I-ACE features support the dietician-patient team in building and tracking an individually tailored healthy lifestyle program.

I-ACE was designed with a Windows (Microsoft Corporation) platform. It has system tables that incorporate food and nutrient databases and evidence-based age-specific, sex-specific, or health status–specific goal packages [3,23-25]. It makes extensive use of embedded graphics, enables the uploading and modification of pictures/infographics/educational materials, and provides graphic presentation of the patient's progress over time. Experienced clinical dieticians (NYZ, MG, and VKS) provided input and feedback on the software's professional content and clinical use features during the development phase. The dieticians in this study, who were the primary users of I-ACE, were computer-literate and routinely used administrative computer programs in their clinical practice.

In addition to the tools for supporting the clinical counseling sessions, I-ACE has administrative-level tools, which (1) can adapt the counseling support tools for use among patients/clients of different (and multiple) ethnic, age, life stage, and/or health status groups, (2) can document all phases of the consultation process for quality control and effectiveness assessment, and (3) has data importing, exporting, and reporting features to support institutional oversight, evaluation, and research activities.

I-ACE enriches the standard approach to dietary counseling in several ways. It uses a food frequency questionnaire (FFQ) and PA questionnaire to systematically document and quantify habitual lifestyle behaviors (for further details on the questionnaires, see the Cultural Adaptation of the Information Technology Tool section) [26,27]. These patient-reported data are summarized and compared with evidence-based food and nutrient intake goals, modeled by the consultant study dieticians on a Mediterranean diet [24,25] and adapted for people with diabetes. These tools are used to individually tailor, focus, and prioritize the educational and counseling processes through identifying lifestyle behaviors that need improvement as targets for education and behavioral change. The counseling component uses Pareto [28] and motivational interviewing [29] principles to set personalized goals, identify the minimal amount of change needed to achieve the maximal impact, simulate changes, and document the patients' willingness to change. Agreed-upon changes are summarized in a take-home report for the patient and followed up in subsequent counseling sessions. These

features are further described along with sample screenshots in Multimedia Appendix 2.

The software includes embedded, modifiable lifestyle educational content based on the published standards of care for medical nutrition therapy for patients with T2DM [3,23] in an easy-to-understand format [30]. It conveys a general and applied understanding of the difference between complex and simple carbohydrates, carbohydrate exchange portions, the glycemic properties of foods, different types of fats (saturated, unsaturated, and trans), optimal sources of dietary fiber and protein, and nutrient-dense versus nutrient-poor foods. The PA content is based upon the World Health Organization's recommendation of at least 150 min/week of moderate-intensity aerobic PA distributed over most days of the week [3,23].

In summary, the software augments and structures the standard approach to lifestyle counseling by providing systematic documentation and quantification of lifestyle data and a graphic interface for education, goal setting, problem solving, and individual tailoring, many aspects of which are modifiable. These features provide new tools to support and expand the active participation of the counselee in making a behavior change plan that suits his/her life.

Cultural Adaptation of the Information Technology Tool

I-ACE provides language congruence through its multilanguage capacity (stage 1: English, Arabic, and Hebrew). In addition, multiple components of the software were culturally adapted for this study. The food database, FFQ, and intervention approach were based on prior epidemiological and interventional research. The I-ACE FFQ included approximately 90 food items that accounted for over 80% of the total energy intake of Arab participants in an earlier, population-based dietary assessment we conducted [26]. In addition, it allowed for other food items to be added from the embedded food database, with 500+ items, developed for the Jewish and Arab populations in Israel [26]. The PA questionnaire was also previously used in our epidemiological research with the Arab population [26,27]. This study intervention was built upon the cultural adaptations that were made for our prior interventional study among Arab women [27]. Those adaptations were directed by the study's Arab coinvestigators and its dieticians, all of whom were from the local Arab community. In addition, focus groups were conducted with local Arab women to obtain their input on cultural and practical aspects of the intervention [27]. For this study, the Arab study investigator and consultant dietician (NYZ) provided input on the adaptation of the software's embedded food photos/graphics and educational materials to reflect the local food customs and cultural norms (see Multimedia Appendix 2 for screen shots of culturally adapted educational materials). The software and cultural content underwent further iterative modifications based on the feedback from the Arab study dieticians. Translation of the infographics and software screens into Arabic was done by a professional translator from the local Arab community and reviewed by the Arab study investigator and dieticians.

Standard Lifestyle Counseling

The comparison SLA arm of the intervention received standard lifestyle counseling as provided by Clalit Health Service dieticians, using existing tools (which did not support quantified dietary assessment) and standard educational materials in Arabic.

Intervention Protocol

Both study arms received 4 individual dietary counseling sessions in the local clinic (in the first, second, third, and sixth months after randomization) and 1 group T2DM self-management session led by a nurse. The first and final dietary counseling sessions (both of which included full dietary assessments) were each approximately 30 to 45 min long for both arms. The median length of the second and third sessions was 17 min each (interquartile range [IQR] 11-26) in the experimental arm, whereas the follow-up visits in the SLA arm were each allotted 15 min, in keeping with the current practice in standard care. Both study arms received very active outreach to encourage adherence to the study protocol and assessments. This differs from standard care, in which intensive, active outreach is not the norm.

We used the same dieticians to provide counseling to both study arms to exclude the possibility of the differences between the groups occurring because of the differences between the dieticians, rather than the intervention type. Most of the intervention was delivered by a single dietician. The dieticians received 2 sessions of 6-hour training for using the I-ACE software before the RCT commenced and had ongoing oversight/support from the principal investigator (KA) and main consultant clinical dietician and coinvestigator (NYZ).

Multimedia Appendix 3 provides a summary of the intervention delivery by study arm.

Intervention adherence in both arms was measured by the attendance of the counseling sessions. The adherence to the use of I-ACE in the dietary counseling sessions in the experimental study arm was measured by checking for the existence of a visit record, including the documented use of the assessment and simulation features.

Outcomes

Primary Outcome

The primary outcome was the diabetes-related dietary knowledge, measured at baseline, 3, 6, and 12 months.

Secondary Outcomes

Secondary outcomes included the (1) dietary intake, measured at baseline and 12 months for all participants (and additionally at 2, 3, and 6 months for participants in the I-ACE arm), (2) leisure PA (LPA), measured at baseline and 12 months, (3) anthropometric measurements (weight and waist circumference [WC]), measured at baseline, 2, 3, 6, and 12 months, and (4) HbA_{1c}, measured before baseline and at 3, 6, and 12 months.

Measures

Primary Outcome (Diabetes Mellitus–Related Dietary Knowledge)

We were not able to find any questionnaires in the literature solely dedicated to DM-related dietary knowledge; however, we did find several general DM knowledge questionnaires that had dietary questions [31-33]. We identified 9 questions from these DM knowledge questionnaires and adapted them to the dietary context of the target population (see examples of the adaptation in Multimedia Appendix 4). We added 2 questions probing the exchange portion sizes and/or the limitations on the intake of certain foods for people with diabetes. Diabetes-related dietary knowledge was assessed as the percentage of correct answers to these questions. An English translation of the DM-related lifestyle knowledge questionnaire is presented in Multimedia Appendix 5. The 2 LPA knowledge questions were taken from existing questionnaires, without need for adaptation [31,34]. The questionnaire was administered at baseline, after the counseling sessions at 3 and 6 months, and at 12 months.

Secondary Outcomes

Dietary Intake

Habitual dietary intake was measured using the computerized I-ACE FFQ, which was based upon FFQs developed by our research group for use among the Arab population in Israel [26,27]. The FFQ was administered to all participants at baseline and 12 months. As the FFQ was used in the experimental arm as a part of the I-ACE counseling approach to track dietary change at each session, the experimental arm also had dietary intake data at 2, 3, and 6 months.

Leisure Physical Activity

Habitual LPA (including the type of activity, frequency, and duration) data were collected using a questionnaire previously used in our research among this population [26,27]. The PA questionnaire was administered at baseline and 12 months.

Anthropometric Measurements

At baseline, weight and height (without shoes, in lightweight clothing) were measured with clinic scales and stadiometers and WC was measured at the midspace between the lowest costal margin and the iliac crest with ergonomic circumference measuring tapes (Seca Medical Measurement Systems and Scales). Weight and WC measurements were repeated at each study visit (2, 3, 6, and 12 months). Weight and height were used to calculate BMI.

Hemoglobin A1c

HbA_{1c} test results were extracted from the electronic medical record for all potential participants before the recruitment to determine eligibility. During the study, the participants were requested to do HbA_{1c} tests every 3 months and all HbA_{1c} test results during the 12-month study period were extracted for the final analysis.

Additional Covariates

Demographic (eg, age, marital status, and education) and health status data were collected at baseline. General DM knowledge was evaluated via the Spoken Knowledge in Low Literacy in Diabetes scale (SKILLD) [34], which was administered at baseline and at 12 months. Information on the pharmacological diabetes management regimen was collected from the electronic medical records at baseline and at 12 months.

Participant/User Satisfaction

We developed a set of questions to elicit the participants' perspectives on the utility of the counseling (in terms of improving their understanding and adherence), satisfaction with the counseling, and for those in the I-ACE arm, satisfaction with the software. We also obtained feedback from the study dietician at the end of the intervention on her experience using the software.

All measures were collected by the study coordinator (face-to-face or by phone), aside from the FFQ, which was administered (face-to-face) by the study dietician.

Analytical Methods

For the primary outcome of DM-related dietary knowledge, the mean (SD) values were reported. A linear mixed regression model for repeated measures was used to evaluate the change in dietary knowledge over time, with an interaction term to determine whether the study groups differed from each other over time. There were missing data in this outcome at the different evaluation periods (34% [17/50] at 3 months, 48% [24/50] at 6 months, and 10% [5/50] at 12 months); however, as there were no missing observations at baseline, we compared the mean baseline DM-related dietary knowledge scores in participants with observed values and those with missing values at each evaluation period. Similar averages imply noninformative missingness (ie, missing at random), supporting the appropriateness of using a linear mixed model for repeated measures with maximum likelihood for this analysis.

For continuous secondary outcomes (eg, food and nutrient intakes, BMI, WC, and HbA_{1c}), mean (SD) values were reported. Multiple linear regression models were used to test for the differences between the study groups at 12 months for changes in lifestyle behaviors. Linear mixed regression models for repeated measures were used to evaluate within-person change over time in dietary intakes in the I-ACE arm as repeated measures from each counseling session were available. WC and BMI, which had repeated measurements in both study arms, were evaluated with linear mixed regression models for repeated measures, with interaction terms for study group by time. For the repeated outcome measure HbA_{1c}, we applied a linear mixed model on the log scale (owing to its non-normal distribution), including a random intercept and potential fixed effects (gender, study arm, age, and baseline DM drug therapy), using the *nlme* package in R. We considered several functions for modeling the effect of time at knots around 6 months as HbA_{1c} was not measured exactly at 6 months, including (1) 2 intervals with the cut-point of 6 months, (2) 2 intervals with the cut-point of 8 months, (3) natural cubic splines with 1 knot, and (4) natural cubic splines with 2 knots. We compared the mean square error of the models resulting from the different functions using cross-validation. The predicted values for a typical subject from each study arm were presented in a graph for each time. The

expected difference in HbA_{1c} from baseline to 6 months and 12 months and their 95% CIs were calculated from the chosen model.

For binary secondary outcomes (eg, any LPA and recommended LPA level), raw count (number, %) was reported. Multiple logistic regression models were used to test for the differences between the study groups in these outcomes at 12 months.

Sex was forced into all multivariable models. Other key variables (age, study town, DM therapy, SKILLD score, and education) were entered into the models and only those with $P<.10$ were retained. Given the small sample size, it was important to keep the models as parsimonious as possible.

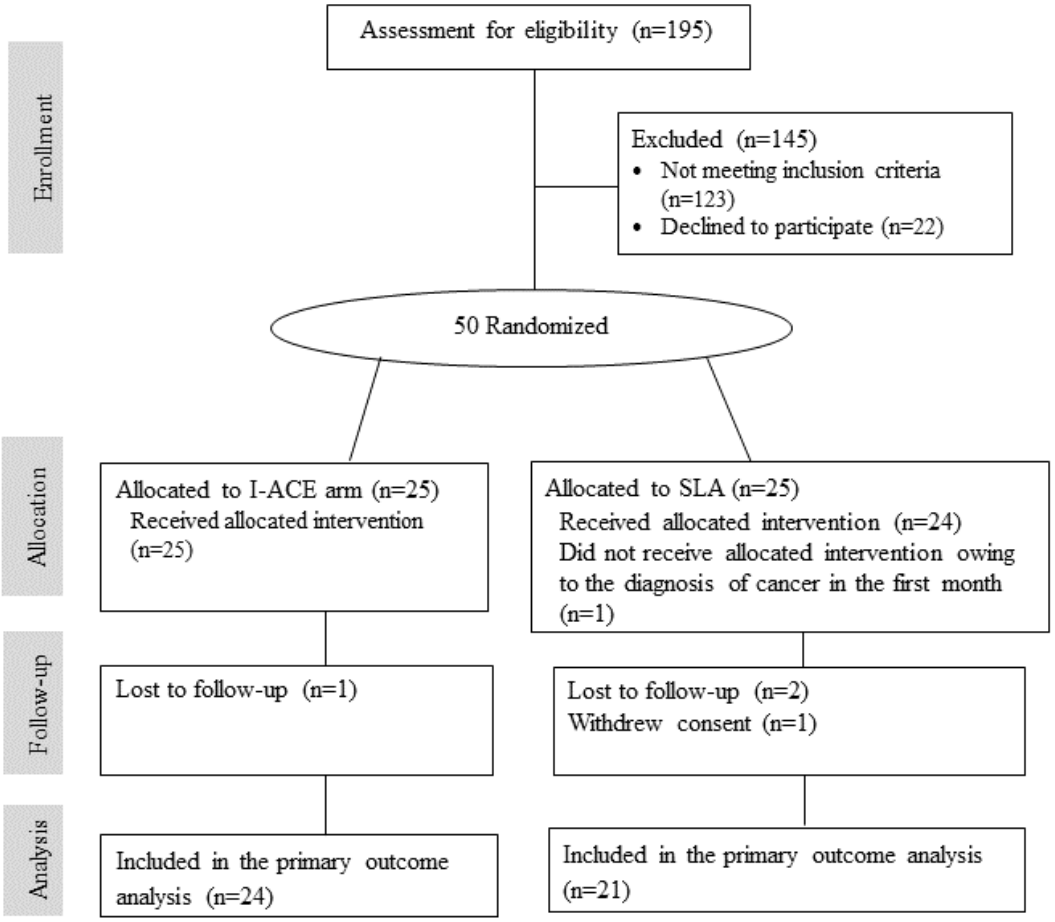
All analyses were performed using SAS version 9.4 SAS Institute, except for the HbA_{1c} repeated measures outcome, which was performed with the open-source statistical software platform R [35]. Simple tabulation and narrative description were used to report the participant and dietician feedbacks on the utility and satisfaction with the counseling and the I-ACE software.

Results

Participant Flow

The participant flow chart is presented in Figure 1. A total of 195 potential participants were identified from Clalit electronic medical records, of whom 123 did not meet the inclusion criteria, 22 refused to participate, and 50 were randomized (25 to the I-ACE arm and 25 to the SLA arm). All those randomized to the I-ACE arm received the allocated intervention (n=25), and all but 1 participant (who was diagnosed with cancer in the first month after randomization) assigned to the SLA arm received the allocated intervention (n=24). One patient with confirmed diabetes and baseline HbA_{1c} of 6.1% was included in the study by mistake and randomized to the I-ACE arm. This patient was included in the data analysis. Furthermore, 3 participants, 1 in the I-ACE arm and 2 in the SLA arm, were lost to follow-up and 1 participant in the SLA arm withdrew consent.

Figure 1. Screening, randomization, and completion of follow-up flow chart for the pilot trial of a culturally-adapted lifestyle counseling software among Arab adults with T2DM. T2DM: Type 2 diabetes mellitus; I-ACE: Interactive lifestyle Assessment, Counseling and Education; SLA: Standard Lifestyle Advice.



Missing Data

Data on the primary outcome (DM-related dietary knowledge) were collected from all 50 participants at baseline (25 in the I-ACE arm and 25 in the SLA arm), but was missing for 17 participants (34%) at the 3-month evaluation (7 in the I-ACE arm and 10 in the SLA arm), 24 participants (48%) at the 6-month evaluation (13 in the I-ACE arm and 11 in the SLA arm), and 5 participants (10%) at the 12-month evaluation (1 from the I-ACE arm and 4 from the SLA arm). To check whether the missing data pattern was informative, we compared the average baseline DM-related dietary knowledge score of participants with observed values to those with missing values at each evaluation period. We did not find significant differences at any of the periods, implying noninformative missingness (see Multimedia Appendix 6).

Information on between-group change in dietary behaviors from baseline to 12 months was missing for 9 participants (4 from the I-ACE arm and 5 from the SLA arm). The analysis of within-group change in dietary behaviors across the intervention counseling visits included all I-ACE participants with more than 1 counseling visit (n=25). The LPA outcomes were missing for 5 participants (1 from the I-ACE arm and 4 from the SLA arm). All 50 participants (25 in each arm) were included in the WC analysis and HbA_{1c} analyses. Counseling satisfaction and utility questionnaires were missing for 5 participants (2 from the I-ACE arm and 3 from the SLA arm). All analyses were conducted according to the originally assigned study groups.

Baseline Data

Table 1 presents the baseline characteristics of the participants, none of which differed significantly by study arm. The average age of the participants at baseline was 53 years, and over 60% were treated with basal insulin, with or without additional oral hypoglycemic therapy. The mean baseline HbA_{1c} was above 9.0% in both study arms.

Table 2 presents the baseline diabetes knowledge and lifestyle behavior data. Participants did not differ on the overall general diabetes knowledge score as measured by the SKILLD scale, which was below 50 on a scale of 100. There were also no significant differences in baseline lifestyle (diet and PA) knowledge or behavior variables by study arm. Participants had suboptimal intakes of dietary fiber, vegetables, and whole grains and very low participation in LPA.

To construct the DM-related lifestyle knowledge score, we excluded the knowledge questions for which ≥85% of participants gave correct answers at baseline, to focus the score on items in need of and amenable to modification by the intervention. The items included in the score are noted in Table 2.

Table 1. Baseline characteristics of 50 Arab participants with type 2 diabetes mellitus in the pilot trial of a culturally adapted lifestyle counseling information technology tool by study group.

Participant characteristics	Total (N=50)	Study arm		*P* value
		I-ACE[a] (n=25)	SLA[b] (n=25)	
Sex (female), n (%)	29 (58)	17 (68)	12 (48)	.15
Age (years), mean (SD)	53.0 (7.6)	52.8 (7.9)	53.2 (7.4)	.87
Married, n (%)	43 (86)	21 (84)	22 (88)	.68
Education (years), mean (SD)	10.8 (3.7)	10.6 (4.2)	11.0 (3.2)	.71
Employed, n (%)	19 (38)	9 (36)	10 (40)	.77
Study town, n (%)				.57
1	26 (48)	12 (48)	14 (56)	—[c]
2	24 (52)	13 (52)	11 (44)	—
Dyslipidemia[d], n (%)	41 (82)	21 (84)	20 (80)	.71
Hypertension[d], n (%)	22 (44)	10 (40)	12 (48)	.57
Cardiovascular disease[d], n (%)	11 (22)	4 (16)	7 (28)	.31
Number of chronic conditions[d], mean (SD)	3.8 (1.5)	3.9 (1.6)	3.7 (1.5)	.64
Physical disability, n (%)	12 (24)	6 (24)	6 (24)	>.99
Age at DM[e] diagnosis (years), mean (SD)	43.8 (7.7)	43.2 (7.6)	44.5 (7.8)	.56
DM duration (years), mean (SD)	9.2 (5.4)	9.6 (5.8)	8.7 (5.0)	.55
DM therapy, n (%)				.88
Diet	1 (2)	0 (0)	1 (4)	—
OHT[f]	18 (36)	10 (40)	8 (32)	—
Basal insulin	2 (4)	1 (4)	1 (4)	—
Basal insulin+OHT	29 (58)	14 (56)	15 (60)	—
Hemoglobin A_{1c} at baseline (%), mean (SD)	9.2 (1.1)	9.1 (1.3)	9.3 (1.0)	.57
Body mass index (kg/m^2), mean (SD)	33.0 (4.1)	33.9 (4.3)	32.1 (3.7)	.17
Waist circumference (cm), mean (SD)	108.2 (9.8)	108.4 (9.9)	108.1 (9.9)	.90
Know last hemoglobin A_{1c} test result, n (%)	39 (78)	21 (84)	18 (72)	.31
Self-blood glucose monitoring frequency, n (%)				.57
Daily	13 (26)	8 (32)	5 (20)	—
Several times a week	12 (24)	6 (24)	6 (24)	—
At least once a month but less than weekly	12 (24)	4 (16)	8 (32)	—
Rarely/never	12 (24)	6 (24)	6 (24)	—

[a]I-ACE: Interactive Lifestyle Assessment, Counseling, and Education.

[b]SLA: standard lifestyle advice.

[c]Not applicable.

[d]On the basis of self-reported physician diagnosis or medical therapy.

[e]DM: diabetes mellitus.

[f]OHT: oral hypoglycemic therapy.

Table 2. Baseline levels of correct diabetes mellitus–related knowledge and lifestyle behaviors among 50 Arab patients with type 2 diabetes mellitus in the Interactive Lifestyle Assessment, Counseling, and Education pilot trial by study group.

Diabetes-related knowledge (% correct) and lifestyle behaviors	Total (N=50)	Study arm		*P* value
		I-ACE[a] (N=25)	SLA[b] (N=25)	
SKILLD[c] general DM[d] knowledge				
Signs of high blood sugar, n (%)	14 (28)	9 (36)	5 (20)	.21
Signs of low blood sugar, n (%)	10 (20)	3 (12)	7 (28)	.16
What to do if blood sugar level is too low, n (%)	4 (8)	0 (0)	4 (16)	.04
Frequency of self-foot check, n (%)	21 (42)	10 (40)	11 (44)	.78
Rationale for self-foot check, n (%)	29 (58)	14 (56)	15 (60)	.78
Frequency and rationale for having eyes checked, n (%)	33 (66)	19 (76)	14 (56)	.14
Normal fasting blood sugar level, n (%)	39 (78)	19 (76)	20 (80)	.73
Normal hemoglobin A_{1c} level, n (%)	30 (60)	16 (64)	14 (56)	.56
Frequency and length of LPA[e] per week, n (%)	18 (36)	10 (40)	8 (32)	.56
Long-term complications of uncontrolled DM, n (%)	44 (88)	22 (88)	22 (88)	>.99
SKILLD score, mean (SD)	48.4 (20.6)	48.8 (19.4)	48.0 (22.2)	.89
DM-related dietary knowledge				
Know limitations[f] on the consumption of:				
Honey[g], n (%)	6 (12)	4 (16)	2 (8)	.38
Dates[g], n (%)	35 (70)	20 (80)	15 (60)	.12
Yogurt/buttermilk[g], n (%)	27 (54)	13 (52)	14 (56)	.78
Cola (regular and nondiet)[g], n (%)	1 (2)	0 (0)	1 (4)	.31
Vegetable salad, n (%)	46 (92)	24 (96)	22 (88)	.30
Rice, n (%)	49 (98)	25 (100)	24 (96)	.31
Pita/bread, n (%)	45 (90)	23 (92)	22 (88)	.64
Cookies (nondiet)[g], n (%)	15 (30)	5 (20)	10 (40)	.12
Grapes[g], n (%)	31 (62)	18 (72)	13 (52)	.15
Fruit juice[g], n (%)	3 (6)	0 (0)	3 (12)	.07
Special (sugar-free) food products for diabetics, n (%)	43 (86)	23 (92)	20 (80)	.22
Identify food highest in carbohydrates[g], n (%)	39 (78)	18 (72)	21 (84)	.31
Identify healthy fat source[g], n (%)	24 (48)	11 (44)	13 (52)	.57
Identify standard exchange portion for, n (%)				
Pita[g]	4 (8)	2 (8)	2 (8)	>.99
Rice[g]	7 (14)	3 (12)	4 (16)	.68
Apple[g]	38 (76)	20 (80)	18 (72)	.51
Yogurt/buttermilk[g]	39 (78)	21 (84)	18 (72)	.31
Dried dates[g]	16 (32)	8 (32)	8 (32)	>.99
Identify food that raises blood sugar the fastest[g], n (%)	31 (62)	17 (68)	14 (56)	.38
Identify food that raises blood sugar most slowly[g], n (%)	20 (40)	11 (44)	9 (36)	.56
Identify best food/drink to treat hypoglycemia[g], n (%)	29 (58)	13 (52)	16 (64)	.39

Diabetes-related knowledge (% correct) and lifestyle behaviors	Total (N=50)	Study arm		*P* value
		I-ACE[a] (N=25)	SLA[b] (N=25)	
Know the effect of physical activity on blood sugar, n (%)	49 (98)	25 (100)	24 (96)	.31
Special (sugar-free) food products not essential to glycemic control[g], n (%)	30 (60)	15 (60)	15 (60)	>.99
Limiting salt intake reduces blood pressure, n (%)	45 (90)	21 (84)	24 (96)	.16
Reducing SFA intake reduces cardiovascular disease risk, n (%)	47 (94)	24 (96)	23 (92)	.55
All questions (% correct answers), mean (SD)	56.7 (9.0)	57.5 (10.2)	55.9 (7.6)	.51
DM-lifestyle knowledge score (% correct answers), mean (SD)	44.4 (10.9)	45.3 (12.1)	43.5 (9.6)	.58
Dietary behaviors, mean (SD)				
Added sugar (% of total energy)	5.4 (0.1)	5.5 (0.1)	5.4 (0.1)	.99
Dietary fiber (g/1000 kcal)	9.6 (2.5)	9.7 (2.6)	9.5 (2.4)	.89
Fruit (portions/day)	3.0 (1.7)	3.0 (1.7)	2.9 (1.7)	.66
Vegetables (portions/day)	3.8 (2.2)	3.4 (1.9)	4.2 (2.4)	.31
Whole grains (portions/day)	2.2 (2.6)	2.2 (2.7)	2.2 (2.5)	.86
LPA[f] behaviors, n (%)				
Any LPA	9 (18)	5 (20)	4 (16)	>.99
≥150 min LPA/week	4 (8)	2 (8)	2 (8)	>.99

[a]I-ACE: Interactive Lifestyle Assessment, Counseling, and Education.

[b]SLA: standard lifestyle advice.

[c]SKILLD: Spoken Knowledge in Low Literacy in Diabetes.

[d]DM: diabetes mellitus.

[e]LPA: leisure physical activity.

[f]From the following categories: completely forbidden, only to be consumed to treat a hypoglycemic episode, can be consumed in limited amount, and can be consumed without limitation.

[g]Items included in the diabetes mellitus–related dietary knowledge score.

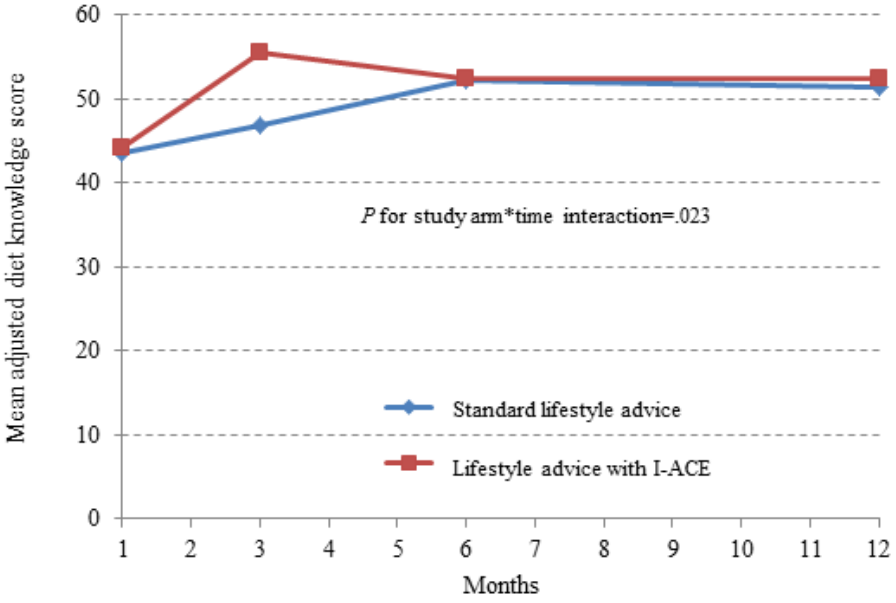
Study Process

The mean (SD) number of counseling sessions attended (out of a maximum of 4) did not differ between the I-ACE and SLA arms (3.44 [SD 0.96] vs 2.92 [SD 1.22] respectively; *P*=.10) Over two-thirds (68%) of the participants in the I-ACE arm attended all 4 counseling sessions and another 16% attended 3 sessions, representing a high adherence of 84%. In the SLA arm, 48% and 16% of the participants attended 4 and 3 counseling sessions, respectively, totaling 64% with high adherence. Session attendance adherence tended to be higher among women than men, but this varied by study arm. Similar proportions of men and women attended at least 3 sessions in the I-ACE arm (88% vs 82%, respectively; *P*=.46); whereas the proportion tended to be lower for men than women in the SLA arm (54% vs 75%, respectively; *P*=.03; see Multimedia Appendix 7 for the number of visits and qualitative information about adherence barriers). The digital visit records saved in the I-ACE software for the experimental arm indicated the intervention adherence (ie, the use of the software for assessment and counseling) in 100% of first visits and 85% of follow-up visits.

Outcomes

The primary study outcome, change in DM-related lifestyle knowledge from baseline, differed between the study arms over time (Figure 2; *P* value for study arm×time interaction=.02). Multimedia Appendix 8 shows the results of the linear mixed model for repeated measures testing the within-group and between-group differences in the knowledge score from baseline to 3, 6, and 12 months. In the I-ACE arm, the mean knowledge score was significantly higher at 3 months than the baseline score and a significant difference was maintained through 12 months. In the SLA arm, the accrual of knowledge occurred more slowly and did not differ significantly from baseline until 6 months. The difference in the slope of the knowledge scores between the study arms was significant at 3 months (higher for I-ACE, *P*=.03). From the sixth month onward, both groups had approximately equal levels of DM-related lifestyle knowledge, which remained significantly higher at 12 months than their mean baseline scores. The increase in knowledge was significantly greater for women than for men (mean [SE] difference=5.14 [SE 2.28]; *P*=.03), whereas, the median (IQR) educational level was lower for women than for men (9 [IQR 8-12] years vs 12 [IQR 10-15] years, respectively; *P*=.049).

Figure 2. Change in DM-related lifestyle knowledge score during intervention (up to 6 months) and follow-up (up to 12 months) among 50 Arab patients with type 2 diabetes mellitus in the I-ACE pilot trial by study arm. Results of a linear mixed regression model for repeated measures with a time*study arm interaction, controlling for sex, educational level, and number of study dietary counseling visits. DM: diabetes mellitus; I-ACE: Interactive lifestyle Assessment.



Multimedia Appendix 9 presents the change in lifestyle behaviors from baseline to 12 months. With regard to the dietary behaviors recommended for diabetes management (eg, adequate consumption of nonstarchy vegetables, whole grains, and dietary fiber and limited consumption of fruit and added sugar), the I-ACE arm exhibited a significant reduction in added sugar intake and a significant increase in dietary fiber intake. The SLA arm exhibited a positive change with regard to a significant decrease in fruit intake and a marginally significant increase in dietary fiber intake but a negative change with regard to a significant decrease in vegetable intake. Although there was a trend toward greater improvement in dietary behaviors in the I-ACE arm than the SLA arm at 12 months, none of the differences reached statistical significance, with the exception of a marginally significant lower intake of added sugar in the I-ACE arm.

As the I-ACE software enabled collecting dietary intake data at each counseling session, we were able to evaluate the within-person change in dietary behaviors across the study visits in the experimental arm (Table 3). Significant changes in the desired direction for all dietary variables occurred during the most intensive phase of the intervention (from baseline to 3 months), during which there were monthly counseling sessions. After that, there was a regression toward baseline intakes and only the changes in added sugar and dietary fiber from baseline remained significant at 12 months.

Table 3. Within-group differences in dietary behaviors from baseline to 2, 3, 6, and 12 months for 25 Arab patients with type 2 diabetes mellitus in the Interactive Lifestyle Assessment, Counseling, and Education pilot trial study arm.

Dietary variable	Difference between intake at baseline and at:[a]							
	2 months		3 months		6 months		12 months	
	Mean (SE)	*P* value	Mean (SE)	*P* value	Mean (SE)	*P* value	Mean (SE)	*P* value
Added sugar (% of total energy)	−1.8 (0.6)	.008	−1.9 (0.7)	.02	−0.9 (0.9)	.34	−2.6 (1.0)	.03
Dietary fiber (g/1000 kcal)	3.7 (0.6)	<.001	3.9 (0.7)	<.001	3.4 (0.8)	<.001	2.7 (0.9)	.003
Fruit (portions/day)	−0.7 (0.2)	.008	−0.7 (0.3)	.048	−0.5 (0.4)	.29	−0.4 (0.4)	.30
Vegetables (portions/day)	1.5 (0.2)	<.001	1.5 (0.3)	<.001	1.0 (0.4)	.02	0.1 (0.4)	.90
Whole grains (portions/day)	2.0 (0.4)	<.001	2.2 (0.5)	<.001	1.2 (0.6)	.09	−0.2 (0.6)	.75

[a]Multivariable linear mixed models for repeated measures controlling for sex. *P* value adjusted for multiple comparisons.

The increase in the percentage of participants in the I-ACE arm reporting any LPA (sex-adjusted odds ratio [OR] 2.80, 95% CI 0.67-11.58; *P*=.16) or the recommended LPA level (sex-adjusted OR 5.01, 95% CI 0.52-47.92; *P*=.16) was greater than that in the SLA arm (Multimedia Appendix 9). However, given the

small sample size, the differences did not reach statistical significance.

For the repeated HbA_{1c} measurements, the model with natural splines with 2 knots (at 5 and 8 months) best fit the data. In addition, the final model included gender. The effect of the study arm was nonsignificant (P=.40); therefore, it was omitted from the model. The effect of time was significant (P<.001). According to the model, the mean HbA_{1c} values were expected to decrease from baseline levels by 11% after the 6-month follow-up (6 months/baseline HbA_{1c} ratio=0.89, 95% CI 0.88-0.91) and by 7% after the 12-month follow-up (12 months/baseline HbA_{1c} ratio=0.93, 95% CI 0.92-0.95). For example, the average HbA_{1c} level at baseline of 9.1% (95% CI 8.8-9.5) was expected to decline to 8.1% (95% CI 8.0%-8.2%) at the 6-month follow-up point and to 8.5% (95% CI 8.4%-8.6%) at the 12-month follow-up point.

Figure 3 presents the predicted HbA_{1c} (in original units) for a typical subject from the SLA and I-ACE study arms. For the purpose of this figure, the study arm was not omitted from the model, although it was nonsignificant. HbA_{1c} levels decreased over time during the active intervention period (1-6 months) for both study arms, reaching a minimum value at approximately 6 months. During the postintervention follow-up period, the levels slightly increased but were still lower at 12 months than baseline levels. There was no significant change in WC over time and this did not differ by study arm in multivariable analysis. Crude results were similar for BMI, and owing to the nonnormal distribution of this variable and its close correlation with WC, multivariable analysis was not conducted for BMI.

Over 90% of the participants in both study arms expressed high satisfaction with the dietician and lifestyle counseling. Most participants indicated that they understood and could implement/utilize the information provided on general nutritional recommendations, portion sizes, exchanges portions, sample menus, and their progress over time; however, the proportion who indicated that they could only understand or implement these materials partially, or not at all, tended to be higher in the SLA than the I-ACE arm (Figure 4).

Figure 3. Expected HbA_{1c} values over time among 50 Arab patients with type 2 diabetes mellitus in the I-ACE pilot trial by study arm. Results from a linear mixed regression model for repeated measures, controlling for sex. HbA_{1c}: hemoglobin A1c; I-ACE: Interactive lifestyle Assessment.

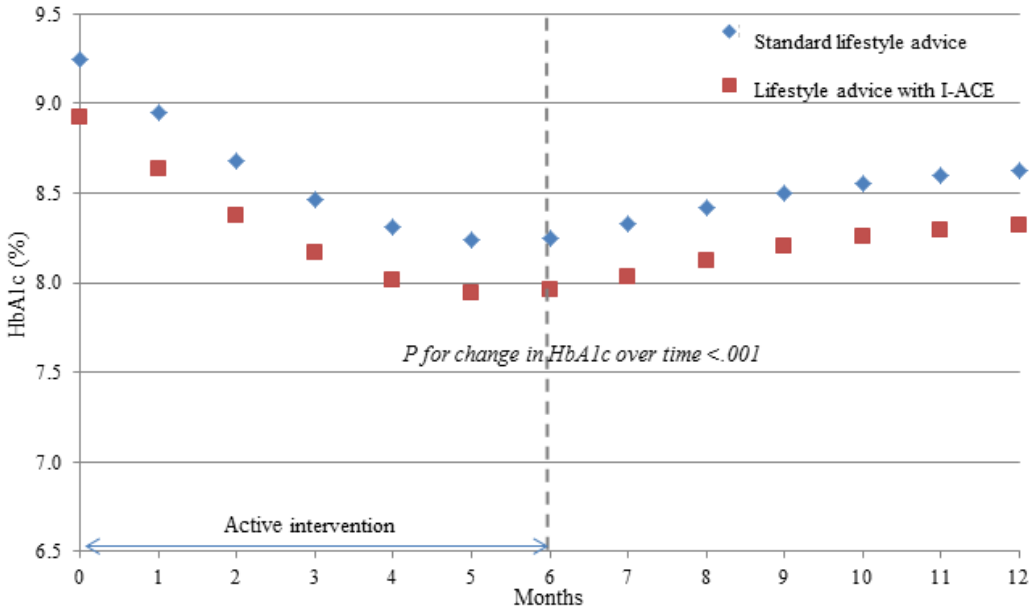
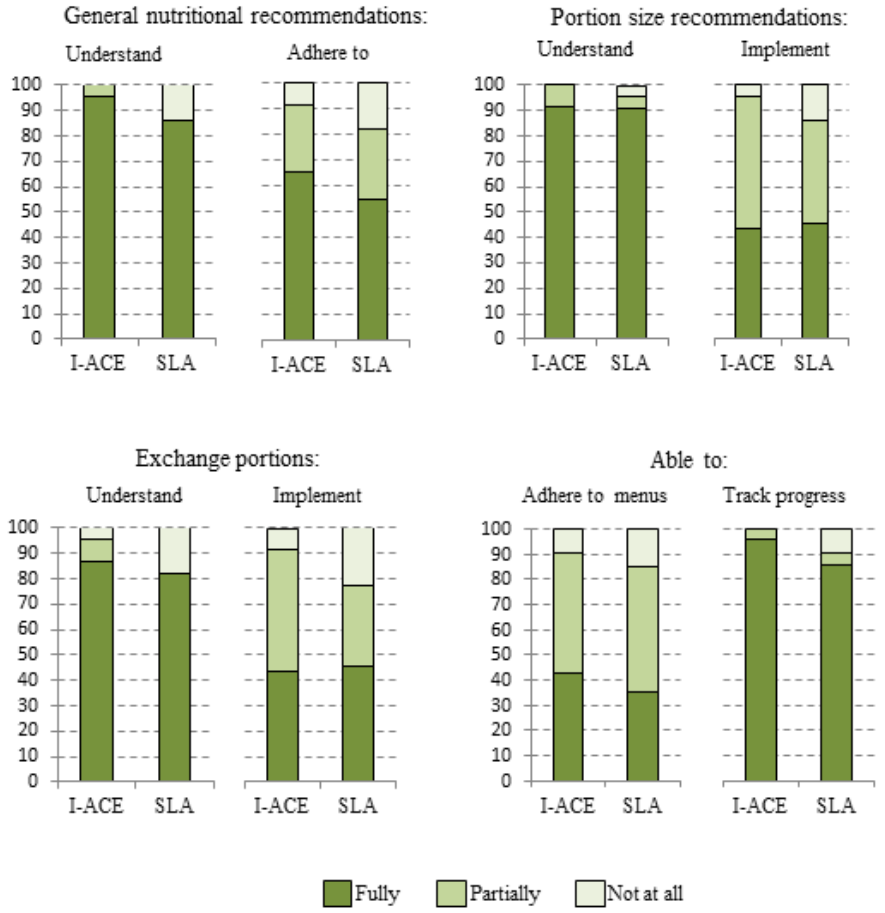


Figure 4. Participant responses to counseling utility questions regarding their ability to understand material and/or adhere to recommendations in the I-ACE pilot trial by study arm. Abbreviations: I-ACE Interactive lifestyle Assessment, Counseling, and Education; SLA Standard Lifestyle Advice.



Most participants in the I-ACE arm thought that the I-ACE software was helpful to the dietary counseling (91%) and did not detract from their interaction with the dietician (96%). The study dietician also expressed overall satisfaction with the I-ACE software. She found the pictorial educational materials and the quantitative nutrient information, provided when individually tailoring sample menus, particularly helpful. She observed that younger participants who were technology oriented engaged more with the software in the counseling sessions than older participants. She noted areas that the software did not address (eg, emotional distress of coping with multiple comorbidities or other personal issues and other emotional support issues that may impact lifestyle behaviors and the readiness to make lifestyle changes).

Harms

We have no harms or unintended consequences to report.

Discussion

Principal Findings

This pilot trial provided indications that the use of the culturally adapted I-ACE software for dietician-delivered lifestyle counseling can increase the pace of acquiring DM-related lifestyle knowledge and showed a trend toward improving lifestyle behaviors (diet and LPA). It was further associated with improved HbA_{1c} results, but not to a greater extent than the dietician-delivered SLA.

This study adds to the body of evidence supporting the efficacy of dietician-delivered dietary counseling for improving dietary knowledge and behaviors and health outcomes in patients with T2DM [3,36]. Consistent with our findings in this study, IT-assisted interventions among T2DM patients have resulted in more rapid knowledge acquisition than non-IT-assisted interventions, particularly when they utilize a combination of provider contact and technology [4,36].

This also holds for ethnic/racial minority groups, particularly when the interventions are culturally adapted/specifically tailored for the target groups [37,38]. In this trial, both groups received written materials in Arabic and dietary counseling from a culturally and linguistically congruent dietician. The digital platform used in the I-ACE arm enabled additional enhancement of cultural adaptation and individual tailoring. The study dietician had access to pictures and infographics reflecting ethnic dietary habits and commonly consumed foods to support the counseling process, both for educational purposes and for creating individually tailored sample menus.

In addition, I-ACE facilitated the use of a broad range of behavioral change modeling techniques to enable active interaction and immediate feedback. These included motivational interviewing, goal setting, goal modification,

simulation of the effects of specific lifestyle changes, and tracking of progress. The Alive! IT-based lifestyle intervention, conducted in a workplace rather than specifically among T2DM patients, employed a similarly broad suite of behavioral change techniques (including the less common component of simulation) and showed significant lifestyle change compared with a wait-listed control group [39,40].

As an additional clinical counseling support feature, I-ACE makes it feasible to collect and update quantifiable data on changes in dietary behaviors at each counseling session, providing a new resource for tracking behavioral change over time. Furthermore, the documentation of the visit record and the dietary counseling process provides a data resource for quality control and other organizational purposes.

The I-ACE tool differs from existing lifestyle IT tools/applications, which typically rely upon multiple administrations of *daily trackers* to collect lifestyle data, resulting in a heavy user burden, user fatigue, and disuse [41]. In addition, these tools/apps are generally used independently by patients, and thus, quantified, actionable information about dietary self-management may not be readily accessible and is not a documented component of the clinical visit record [42-48]. Numerous studies have shown lifestyle-related IT tools to be more effective when used in clinical settings with health care provider contact than when used independently [4,36,49].

Few apps developed specifically for T2DM self-management include the assessment of dietary/lifestyle habits. Those that have included this domain either (1) did so in a superficial and nonquantified manner and did not improve dietary behaviors [44] or (2) required detailed recording of daily dietary intake, which resulted in a very high user burden [50] and was thus one of the least used features of mobile T2DM support apps [45]. Other apps/Web-based IT tools included unidirectional messages from the app to the patient or noninteractive Web-based information that did not support any individual tailoring of dietary/lifestyle plans [51,52].

Users' satisfaction with I-ACE was high, suggesting it may enhance the dietary counseling process. Responses of the I-ACE arm participants to the utility questions also suggest that it may have potential for optimizing the patient's understanding of and adherence to general nutritional recommendations, portion size and exchange portion recommendations, individually tailored menus, and the tracking of progress over time.

Reviews of the literature show that dietary counseling, particularly among patients with T2DM, improves the glycemic control to the same magnitude as expected from introducing a new drug [3,36-38,53]. In this pilot trial, we found that a short, dietician-delivered dietary counseling series (eg, 4 sessions over the course of 6 months) produced a decrease of 1% point from baseline HbA_{1c} at the end of the active intervention period and a 0.5% decrease at 12 months (6 months after the active intervention ended), in a sample starting with poor glycemic control.

At the same time, the limited number of dietary consultations in this pilot trial did not appear to be adequate to ensure fully sustained improvements in dietary behaviors and glycemic control. Systematic reviews of diabetes self-management education and nutritional interventions among adults with T2DM have shown a higher success rate in interventions with over 10 hours of contact with providers [4]. Interventions that provided dietary counseling encounters for more than 6 months reported that the improvement and continued reduction of HbA_{1c} was maintained for up to 2 years [3].

This pilot study has several strengths and limitations. The study adds to the very limited literature on culturally adapting and testing IT tools for low-SES ethnic minority populations and suggests that such tools can improve the efficiency of DM-related dietary education. The RCT design adds also to the limited body of RCTs evaluating the IT-supported lifestyle and educational interventions in low-SES minority communities, which typically have disproportionately high T2DM rates and few well-designed studies targeting the key self-management domain of nutrition among T2DM patients.

Although the study was adequately powered to test the primary outcome of DM-related dietary knowledge, it had limited power to test the secondary outcomes. Nevertheless, we observed trends toward greater improvement in lifestyle behaviors in the I-ACE arm than the SLA arm. This is particularly impressive as, from the outset, we aimed to recruit a challenging group of patients who were overweight or obese and had poor glycemic control, suggesting that despite having free access to health care services and subsidized medications, their engagement with and benefit from existing health services was suboptimal. Our experience with the intervention indicated that adequate resources and maximum institutional flexibility and accessibility are needed to insure successful implementation, particularly in vulnerable, low-SES populations. Nevertheless, we observed reasonably good adherence to the study intervention, with a suggested trend of improved engagement among men in the I-ACE arm so that it equaled that of women; whereas in the SLA arm, the men's engagement was lower than that of women. Given the small sample in our pilot RCT, these suggested trends should be confirmed in a larger efficacy trial.

We used the same dieticians to deliver dietary counseling in both study arms to eliminate the possibility of differences between study arms because of different dieticians. The use of the I-ACE counseling approach may have affected the way the dietician delivered standard care in the comparison arm. In addition, the study was conducted in community clinics serving local neighborhoods comprised extended families. As the randomization was conducted on an individual basis, neighbors and extended family members who were assigned to different study arms may have had shared what they learned from the experimental intervention approach and materials outside of the clinic setting. Both these possible sources of contamination, however, would have led to an underestimation of the I-ACE experimental effect and thus do not detract from the significance of our findings. We recommend that a definitive RCT should consider a cluster randomized design, particularly if conducted in neighborhood clinics with very strong, extensive social networks.

Finally, the intervention in this pilot study was of limited duration and the findings suggest that a longer/more intensive intervention is important to assure sustainability.

Generalizability

This study was conducted in a low-SES ethnic minority community at high risk of T2DM and its complications and showed potential for improving self-care, promoting healthy lifestyle behaviors, and improving surrogate health outcomes (HbA_{1c}). The findings may be informative for planning definitive RCTs to evaluate IT-based clinical interventions in other similar populations.

The use of the I-ACE software took more time than is currently allocated in clinical care in Israel, particularly for the initial dietary counseling session (35-45 min), although the median time for follow-up visits (17 min) did not differ substantially from the 15-min slots allotted in routine follow-up care. This time differential, which is largely the result of the structured task flow, the additional tools at the dietician's disposal, and the associated documentation, needs to be taken into account when considering application to clinical care. One possible accommodation would be to model routine clinical care differently. For example, the national health insurance basket of services in Israel includes 14 visits per year to a dietician for patients with DM, which, in practice, are rarely fully utilized. An alternative model worth testing would be the integration of I-ACE in a series of fewer but longer dietary counseling sessions, with a comparison of its effectiveness with the usual care with multiple short visits. In addition, further technological development, such as making the assessment and other features of the I-ACE software available for self-use by patients via mobile or Web apps, could reduce the time required for documentation in the face-to-face consultations with the dietician.

Conclusions

This pilot trial supports the potential of the dietician-delivered, culturally adapted I-ACE dietary counseling intervention to increase the efficiency of DM-related lifestyle education and to improve lifestyle behaviors, compared with the usual dietary counseling in a minority population, although confirming that both types of counseling, coupled with active outreach, improved glycemic control. These findings provide support for conducting a large-scale trial to evaluate the effectiveness and applicability of the I-ACE software in routine clinical care among ethnically diverse populations.

Acknowledgments

The authors would like to thank the study participants, dieticians, and coordinators. The authors also acknowledge the important contributions and assistance of the clinics' physicians and administrative staff.

This study was supported by a grant from the European Foundation for the Study of Diabetes/European Diabetes Research Programme in Patient Education from AstraZeneca/Bristol-Myers Squibb. The study sponsor approved the study design but had no role in the conduct of the study; in the collection, analysis, and interpretation of the data; or in the preparation, review, or approval of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
CONSORT‐EHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 16146 KB - jmir_v21i10e13674_app1.pdf]

Multimedia Appendix 2
Selected I-ACE screen shots.
[PDF File (Adobe PDF File), 987 KB - jmir_v21i10e13674_app2.pdf]

Multimedia Appendix 3
Summary of the intervention delivery by study arm.
[PDF File (Adobe PDF File), 10 KB - jmir_v21i10e13674_app3.pdf]

Multimedia Appendix 4
Examples of modification of diet knowledge questions.
[PDF File (Adobe PDF File), 72 KB - jmir_v21i10e13674_app4.pdf]

Multimedia Appendix 5
DM-related lifestyle knowledge questionnaire. DM: diabetes mellitus.
[PDF File (Adobe PDF File), 137 KB - jmir_v21i10e13674_app5.pdf]

Multimedia Appendix 6
Baseline DM diet-related knowledge scores among participants with and without data at 3, 6 and 12 months. DM: diabetes mellitus.
[PDF File (Adobe PDF File), 10 KB - jmir_v21i10e13674_app6.pdf]

Multimedia Appendix 7
Number of counselling visits by study arm and qualitative information about intervention adherence.
[PDF File (Adobe PDF File), 88 KB - jmir_v21i10e13674_app7.pdf]

Multimedia Appendix 8
Within-group and between-group differences in DM diet-related knowledge from baseline at 3, 6, and 12 months. DM: diabetes mellitus.
[PDF File (Adobe PDF File), 12 KB - jmir_v21i10e13674_app8.pdf]

Multimedia Appendix 9
Difference in lifestyle behaviors at 12 months by study arm.
[PDF File (Adobe PDF File), 88 KB - jmir_v21i10e13674_app9.pdf]

References

1. American Diabetes Association. Standards of medical care in diabetes-2018. Diabetes Care 2018 Jan;41(Suppl 1):S1-S2. [doi: 10.2337/dc18-Sint01] [Medline: 29222369]
2. American Diabetes Association. Improving care and promoting health in populations: standards of medical care in diabetes-2018. Diabetes Care 2018 Jan;41(Suppl 1):S7-12. [doi: 10.2337/dc18-S001] [Medline: 29222372]
3. Franz MJ, MacLeod J, Evert A, Brown C, Gradwell E, Handu D, et al. Academy of nutrition and dietetics nutrition practice guideline for type 1 and type 2 diabetes in adults: systematic review of evidence for medical nutrition therapy effectiveness and recommendations for integration into the nutrition care process. J Acad Nutr Diet 2017 Oct;117(10):1659-1679. [doi: 10.1016/j.jand.2017.03.022] [Medline: 28533169]
4. Chrvala CA, Sherr D, Lipman RD. Diabetes self-management education for adults with type 2 diabetes mellitus: a systematic review of the effect on glycemic control. Patient Educ Couns 2016 Jun;99(6):926-943 [FREE Full text] [doi: 10.1016/j.pec.2015.11.003] [Medline: 26658704]
5. Duke SS, Colagiuri S, Colagiuri R. Individual patient education for people with type 2 diabetes mellitus. Cochrane Database Syst Rev 2009 Jan 21(1):CD005268 [FREE Full text] [doi: 10.1002/14651858.CD005268.pub2] [Medline: 19160249]
6. Creamer J, Attridge M, Ramsden M, Cannings-John R, Hawthorne K. Culturally appropriate health education for type 2 diabetes in ethnic minority groups: an updated Cochrane review of randomized controlled trials. Diabet Med 2016 Mar;33(2):169-183. [doi: 10.1111/dme.12865] [Medline: 26202820]
7. Pal K, Eastwood SV, Michie S, Farmer AJ, Barnard ML, Peacock R, et al. Computer-based diabetes self-management interventions for adults with type 2 diabetes mellitus. Cochrane Database Syst Rev 2013 Mar 28(3):CD008776 [FREE Full text] [doi: 10.1002/14651858.CD008776.pub2] [Medline: 23543567]
8. Neuenschwander LM, Abbott A, Mobley AR. Comparison of a web-based vs in-person nutrition education program for low-income adults. J Acad Nutr Diet 2013 Jan;113(1):120-126. [doi: 10.1016/j.jand.2012.07.034] [Medline: 23092741]
9. Prabhakaran D, Jha D, Prieto-Merino D, Roy A, Singh K, Ajay VS, et al. Effectiveness of an mhealth-based electronic decision support system for integrated management of chronic conditions in primary care: the mwellcare cluster-randomized controlled trialEffectiveness of an mhealth-based electronic decision support system for integrated management of chronic conditions in primary care: the m-wellcare cluster-randomized controlled trial. Circulation 2019;139(3):380-391. [doi: 10.1161/CIRCULATIONAHA.118.038192] [Medline: 30586732]
10. Bowen ME, Cavanaugh KL, Wolff K, Davis D, Gregory RP, Shintani A, et al. The diabetes nutrition education study randomized controlled trial: a comparative effectiveness study of approaches to nutrition in diabetes self-management education. Patient Educ Couns 2016 Aug;99(8):1368-1376 [FREE Full text] [doi: 10.1016/j.pec.2016.03.017] [Medline: 27026388]
11. Kattelmann KK, Conti K, Ren C. The medicine wheel nutrition intervention: a diabetes education study with the Cheyenne River Sioux tribe. J Am Diet Assoc 2009 Sep;109(9):1532-1539 [FREE Full text] [doi: 10.1016/j.jada.2009.06.362] [Medline: 19699832]
12. Adam L, O'Connor C, Garcia AC. Evaluating the impact of diabetes self-management education methods on knowledge, attitudes and behaviours of adult patients with type 2 diabetes mellitus. Can J Diabetes 2018 Oct;42(5):470-7.e2. [doi: 10.1016/j.jcjd.2017.11.003] [Medline: 29449096]
13. Daghio MM, Fattori G, Ciardullo AV. Use of pictorial advice to promote compliance to diet and drugs among illiterate and migrant patients. J Diabetology 2010;3:3 [FREE Full text]

14. Misra R, Lager J. Ethnic and gender differences in psychosocial factors, glycemic control, and quality of life among adult type 2 diabetic patients. J Diabetes Complications 2009;23(1):54-64. [doi: 10.1016/j.jdiacomp.2007.11.003] [Medline: 18413181]
15. Isaacs T, Hunt D, Ward D, Rooshenas L, Edwards L. The inclusion of ethnic minority patients and the role of language in telehealth trials for type 2 diabetes: a systematic review. J Med Internet Res 2016 Sep 26;18(9):e256 [FREE Full text] [doi: 10.2196/jmir.6374] [Medline: 27670360]
16. Franz MJ, MacLeod J. Success of nutrition-therapy interventions in persons with type 2 diabetes: challenges and future directions. Diabetes Metab Syndr Obes 2018;11:265-270 [FREE Full text] [doi: 10.2147/DMSO.S141952] [Medline: 29928137]
17. Daoud N, Soskolne V, Mindell JS, Roth MA, Manor O. Ethnic inequalities in health between Arabs and Jews in Israel: the relative contribution of individual-level factors and the living environment. Int J Public Health 2018 Apr;63(3):313-323. [doi: 10.1007/s00038-017-1065-3] [Medline: 29273838]
18. Kalter-Leibovici O, Chetrit A, Lubin F, Atamna A, Alpert G, Ziv A, et al. Adult-onset diabetes among Arabs and Jews in Israel: a population-based study. Diabet Med 2012 Jun;29(6):748-754. [doi: 10.1111/j.1464-5491.2011.03516.x] [Medline: 22050554]
19. Wilf-Miron R, Peled R, Yaari E, Shem-Tov O, Weinner VA, Porath A, et al. Disparities in diabetes care: role of the patient's socio-demographic characteristics. BMC Public Health 2010 Nov 25;10:729 [FREE Full text] [doi: 10.1186/1471-2458-10-729] [Medline: 21108780]
20. Jaffe A, Giveon S, Wulffhart L, Oberman B, Baidousi M, Ziv A, et al. Adult Arabs have higher risk for diabetes mellitus than Jews in Israel. PLoS One 2017;12(5):e0176661 [FREE Full text] [doi: 10.1371/journal.pone.0176661] [Medline: 28481942]
21. Howard-Pitney B, Winkleby MA, Albright CL, Bruce B, Fortmann SP. The Stanford Nutrition Action Program: a dietary fat intervention for low-literacy adults. Am J Public Health 1997 Dec;87(12):1971-1976. [doi: 10.2105/ajph.87.12.1971] [Medline: 9431286]
22. Hawthorne K, Tomlinson S. One-to-one teaching with pictures--flashcard health education for British Asians with diabetes. Br J Gen Pract 1997 May;47(418):301-304 [FREE Full text] [Medline: 9219407]
23. Inzucchi SE, Bergenstal RM, Buse JB, Diamant M, Ferrannini E, Nauck M, American Diabetes Association (ADA), European Association for the Study of Diabetes (EASD). Management of hyperglycemia in type 2 diabetes: a patient-centered approach: position statement of the American Diabetes Association (ADA) and the European Association for the Study of Diabetes (EASD). Diabetes Care 2012 Jun;35(6):1364-1379 [FREE Full text] [doi: 10.2337/dc12-0413] [Medline: 22517736]
24. Schröder H, Fitó M, Estruch R, Martínez-González MA, Corella D, Salas-Salvadó J, et al. A short screener is valid for assessing Mediterranean diet adherence among older Spanish men and women. J Nutr 2011 Jun;141(6):1140-1145. [doi: 10.3945/jn.110.135566] [Medline: 21508208]
25. Hodge AM, English DR, Itsiopoulos C, O'Dea K, Giles GG. Does a Mediterranean diet reduce the mortality risk associated with diabetes: evidence from the Melbourne collaborative cohort study. Nutr Metab Cardiovasc Dis 2011 Sep;21(9):733-739. [doi: 10.1016/j.numecd.2010.10.014] [Medline: 21194913]
26. Abu-Saad K, Murad H, Lubin F, Freedman LS, Ziv A, Alpert G, et al. Jews and Arabs in the same region in Israel exhibit major differences in dietary patterns. J Nutr 2012 Dec;142(12):2175-2181. [doi: 10.3945/jn.112.166611] [Medline: 23096004]
27. Kalter-Leibovici O, Younis-Zeidan N, Atamna A, Lubin F, Alpert G, Chetrit A, et al. Lifestyle intervention in obese Arab women: a randomized controlled trial. Arch Intern Med 2010 Jun 14;170(11):970-976. [doi: 10.1001/archinternmed.2010.103] [Medline: 20548010]
28. Koch R. The 80/20 Principle: The Secret to Achieving More with Less. Third Edition. New York: Doubleday; 2008.
29. Resnicow K, Jackson A, Wang T, De AK, McCarty F, Dudley WN, et al. A motivational interviewing intervention to increase fruit and vegetable intake through black churches: results of the Eat for Life trial. Am J Public Health 2001 Oct;91(10):1686-1693. [doi: 10.2105/ajph.91.10.1686] [Medline: 11574336]
30. Learning About Diabetes. 2019. URL: https://www.learningaboutdiabetes.org/ [accessed 2019-02-03]
31. Michigan Diabetes Research Center (MDRC). Michigan Diabetes Research and Training Center's Revised Diabetes Knowledge Test URL: http://diabetesresearch.med.umich.edu/peripherals/profs/documents/svi/DKT2_with_answers.pdf [accessed 2019-02-03]
32. Garcia AA, Villagomez ET, Brown SA, Kouzekanani K, Hanis CL. The Starr County Diabetes Education Study: development of the Spanish-language diabetes knowledge questionnaire. Diabetes Care 2001 Jan;24(1):16-21. [doi: 10.2337/diacare.24.1.16] [Medline: 11194219]
33. Eigenmann C, Skinner T, Colagiuri R. Development and validation of a diabetes knowledge questionnaire. Pract Diab Int 2011 May 23;28(4):166-70d [FREE Full text] [doi: 10.1002/pdi.1586]
34. Rothman RL, Malone R, Bryant B, Wolfe C, Padgett P, DeWalt DA, et al. The spoken knowledge in low literacy in diabetes scale: a diabetes knowledge scale for vulnerable patients. Diabetes Educ 2005;31(2):215-224. [doi: 10.1177/0145721705275002] [Medline: 15797850]
35. R: The R Project for Statistical Computing. 2017. URL: https://www.r-project.org/ [accessed 2019-02-03]

36. Mitchell LJ, Ball LE, Ross LJ, Barnes KA, Williams LT. Effectiveness of dietetic consultations in primary health care: a systematic review of randomized controlled trials. J Acad Nutr Diet 2017 Dec;117(12):1941-1962. [doi: 10.1016/j.jand.2017.06.364] [Medline: 28826840]
37. Attridge M, Creamer J, Ramsden M, Cannings-John R, Hawthorne K. Culturally appropriate health education for people in ethnic minority groups with type 2 diabetes mellitus. Cochrane Database Syst Rev 2014 Sep 4(9):CD006424. [doi: 10.1002/14651858.CD006424.pub3] [Medline: 25188210]
38. Ricci-Cabello I, Ruiz-Pérez I, Rojas-García A, Pastor G, Rodríguez-Barranco M, Gonçalves DC. Characteristics and effectiveness of diabetes self-management educational programs targeted to racial/ethnic minority groups: a systematic review, meta-analysis and meta-regression. BMC Endocr Disord 2014 Jul 19;14:60 [FREE Full text] [doi: 10.1186/1472-6823-14-60] [Medline: 25037577]
39. Block G, Sternfeld B, Block CH, Block TJ, Norris J, Hopkins D, et al. Development of Alive! (a lifestyle intervention via email), and its effect on health-related quality of life, presenteeism, and other behavioral outcomes: randomized controlled trial. J Med Internet Res 2008 Nov 19;10(4):e43 [FREE Full text] [doi: 10.2196/jmir.1112] [Medline: 19019818]
40. Sternfeld B, Block C, Quesenberry Jr CP, Block TJ, Husson G, Norris JC, et al. Improving diet and physical activity with ALIVE: a worksite randomized trial. Am J Prev Med 2009 Jun;36(6):475-483. [doi: 10.1016/j.amepre.2009.01.036] [Medline: 19460655]
41. Krebs P, Duncan DT. Health app use among US mobile phone owners: a national survey. JMIR Mhealth Uhealth 2015 Nov 4;3(4):e101 [FREE Full text] [doi: 10.2196/mhealth.4924] [Medline: 26537656]
42. Michie S, Yardley L, West R, Patrick K, Greaves F. Developing and evaluating digital interventions to promote behavior change in health and health care: recommendations resulting from an international workshop. J Med Internet Res 2017 Jun 29;19(6):e232 [FREE Full text] [doi: 10.2196/jmir.7126] [Medline: 28663162]
43. Iribarren SJ, Cato K, Falzon L, Stone PW. What is the economic evidence for mhealth? A systematic review of economic evaluations of mhealth solutions. PLoS One 2017;12(2):e0170581 [FREE Full text] [doi: 10.1371/journal.pone.0170581] [Medline: 28152012]
44. Holmen H, Torbjørnsen A, Wahl AK, Jenum AK, Småstuen MC, Arsand E, et al. A mobile health intervention for self-management and lifestyle change for persons with type 2 diabetes, part 2: one-year results from the Norwegian randomized controlled trial RENEWING HEALTH. JMIR Mhealth Uhealth 2014 Dec 11;2(4):e57 [FREE Full text] [doi: 10.2196/mhealth.3882] [Medline: 25499872]
45. Lamprinos I, Demski H, Mantwill S, Kabak Y, Hildebrand C, Ploessnig M. Modular ICT-based patient empowerment framework for self-management of diabetes: design perspectives and validation results. Int J Med Inform 2016 Jul;91:31-43. [doi: 10.1016/j.ijmedinf.2016.04.006] [Medline: 27185507]
46. Ribu L, Holmen H, Torbjørnsen A, Wahl AK, Grøttland A, Småstuen MC, et al. Low-intensity self-management intervention for persons with type 2 diabetes using a mobile phone-based diabetes diary, with and without health counseling and motivational interviewing: protocol for a randomized controlled trial. JMIR Res Protoc 2013 Aug 26;2(2):e34 [FREE Full text] [doi: 10.2196/resprot.2768] [Medline: 23978690]
47. Goyal S, Lewis G, Yu C, Rotondi M, Seto E, Cafazzo JA. Evaluation of a behavioral mobile phone app intervention for the self-management of type 2 diabetes: randomized controlled trial protocol. JMIR Res Protoc 2016 Aug 19;5(3):e174 [FREE Full text] [doi: 10.2196/resprot.5959] [Medline: 27542325]
48. Kitsiou S, Paré G, Jaana M, Gerber B. Effectiveness of mhealth interventions for patients with diabetes: an overview of systematic reviews. PLoS One 2017;12(3):e0173160 [FREE Full text] [doi: 10.1371/journal.pone.0173160] [Medline: 28249025]
49. Almeida VA, Littlejohn P, Cop I, Brown E, Afroze R, Davison KM. Comparison of nutrigenomics technology interface tools for consumers and health professionals: a sequential explanatory mixed methods investigation. J Med Internet Res 2019 Jun 28;21(6):e12580 [FREE Full text] [doi: 10.2196/12580] [Medline: 31254340]
50. Rabbi M, Pfammatter A, Zhang M, Spring B, Choudhury T. Automated personalized feedback for physical activity and dietary behavior change with mobile phones: a randomized controlled trial on adults. JMIR Mhealth Uhealth 2015 May 14;3(2):e42 [FREE Full text] [doi: 10.2196/mhealth.4160] [Medline: 25977197]
51. Murray E, Ross J, Pal K, Li J, Dack C, Stevenson F, et al. National Institute for Health Research. 2018. A Web-Based Self-Management Programme for People With Type 2 Diabetes: The HeLP-Diabetes Research Programme Including RCT URL: https://njl-admin.nihr.ac.uk/document/download/2021255 [accessed 2019-08-29]
52. Boels AM, Rutten G, Zuithoff N, de Wit A, Vos R. Effectiveness of diabetes self-management education via a smartphone application in insulin treated type 2 diabetes patients - design of a randomised controlled trial ('TRIGGER study'). BMC Endocr Disord 2018 Oct 22;18(1):74 [FREE Full text] [doi: 10.1186/s12902-018-0304-9] [Medline: 30348142]
53. Ramadas A, Chan CK, Oldenburg B, Hussein Z, Quek KF. Randomised-controlled trial of a web-based dietary intervention for patients with type 2 diabetes: changes in health cognitions and glycemic control. BMC Public Health 2018 Jun 8;18(1):716 [FREE Full text] [doi: 10.1186/s12889-018-5640-1] [Medline: 29884161]

Abbreviations

BMI: body mass index
DM: diabetes mellitus
FFQ: food frequency questionnaire
HbA_{1c}: hemoglobin A1c
I-ACE: Interactive Lifestyle Assessment, Counseling, and Education
IQR: interquartile range
IT: information technology
LPA: leisure physical activity
OR: odds ratio
PA: physical activity
RCT: randomized controlled trial
SES: socioeconomic status
SKILLD: Spoken Knowledge in Low Literacy in Diabetes
SLA: standard lifestyle advice
T2DM: type 2 diabetes mellitus
WC: waist circumference

Edited by G Eysenbach; submitted 11.02.19; peer-reviewed by K Smith, A Willis, H Lin; comments to author 16.05.19; revised version received 10.07.19; accepted 19.08.19; published 16.10.19.

Please cite as:
Abu-Saad K, Murad H, Barid R, Olmer L, Ziv A, Younis-Zeidan N, Kaufman-Shriqui V, Gillon-Keren M, Rigler S, Berchenko Y, Kalter-Leibovici O
Development and Efficacy of an Electronic, Culturally Adapted Lifestyle Counseling Tool for Improving Diabetes-Related Dietary Knowledge: Randomized Controlled Trial Among Ethnic Minority Adults With Type 2 Diabetes Mellitus
J Med Internet Res 2019;21(10):e13674
URL: https://www.jmir.org/2019/10/e13674
doi:10.2196/13674
PMID:31621640

Review

Identifying Frameworks for Validation and Monitoring of Consensual Behavioral Intervention Technologies: Narrative Review

François Carbonnel[1,2,3,4*], MD, PhD; Gregory Ninot[1,2,5*], PhD

[1]Research Unit EA4556 Epsylon, University of Montpellier, Montpellier, France

[2]Plateforme Universitaire Collaborative d'Evaluation des Programmes de Prévention et de Soins de Support, University of Montpellier, Montpellier, France

[3]University Department of General Practice, University of Montpellier, Montpellier, France

[4]University Multiprofessional Health Center Avicenne, Perpignan, France

[5]Montpellier Cancer Institute, Montpellier, France

[*]all authors contributed equally

Corresponding Author:
François Carbonnel, MD, PhD
University Department of General Practice
University of Montpellier
641 Avenue du Doyen Giraud
UFR Médecine site Nord
Montpellier, 34093
France
Phone: 33 684014834
Email: francois.carbonnel@gmail.com

Abstract

Background: Changing health behaviors, such as smoking, unhealthy eating, inactivity, and alcohol abuse, may have a greater impact on population health than any curative strategy. One of the suggested strategies is the use of behavioral intervention technologies (BITs). They open up new opportunities in the area of prevention and therapy and have begun to show benefits in the durable change of health behaviors in patients or those at risk. A consensual and international paradigm was adopted by health authorities for drugs 50 years ago. It guides their development from research units to their authorization and surveillance. BITs' generalization brings into question their upstream evaluation before being placed on the market and their downstream monitoring once on the market; this is especially the case in view of the marketing information provided by manufacturers and the scarcity and methodological limits of scientific studies on these tools.

Objective: This study aims to identify and categorize the frameworks for the validation and monitoring of BITs proposed in the literature.

Methods: We conducted a narrative literature review using MEDLINE, PsycINFO, and Web of Science. The review items included the following: name, publication year, name of the creator (ie, first author), country, funding organization, health focus, target group, and design (ie, linear, iterative, evolutive, and/or concurrent). The frameworks were then categorized based on (1) translational research thanks to a continuum of steps and (2) the three paradigms that may have inspired the frameworks: biomedical, engineering, and/or behavioral.

Results: We identified 46 frameworks besides the classic US Food and Drug Administration (FDA) five-phase drug development model. A total of 57% (26/46) of frameworks were created in the 2010s and 61% (28/46) involved the final user in an early and systematic way. A total of 4% (2/46) of frameworks had a linear-only sequence of their phases, 37% (17/46) had a linear and iterative structure, 33% (15/46) added an evolutive structure, and 24% (11/46) were associated with a parallel process. Only 12 out of 46 (26%) frameworks covered the continuum of steps and 12 (26%) relied on the three paradigms.

Conclusions: To date, 46 frameworks of BIT validation and surveillance coexist, besides the classic FDA five-phase drug development model, without the predominance of one of them or convergence in a consensual model. Their number has increased exponentially in the last three decades. Three dangerous scenarios are possible: (1) anarchic continuous development of BITs that depend on companies amalgamating health benefits and usability (ie, user experience, data security, and ergonomics) and

limiting implementation to several countries; (2) the movement toward the type of framework for drug evaluation centered on establishing its effectiveness before marketing authorization to guarantee its safety for users, which is heavy and costly; and (3) the implementation of a framework reliant on big data analysis based on a posteriori research and an autoregulation of a market, but that does not address the safety risk for the health user, as the market will not regulate safety or efficacy issues. This paper recommends convergence toward an international validation and surveillance framework based on the specificities of BITs, not equivalent to medical devices, to guarantee their effectiveness and safety for users.

(J Med Internet Res 2019;21(10):e13606) doi:10.2196/13606

KEYWORDS

behavioral intervention technology; validation; surveillance; paradigm; framework; nonpharmacological interventions

Introduction

According to the World Health Organization (WHO), changing health-related behaviors, such as smoking, unhealthy diet, physical inactivity, and alcohol abuse, could avoid up to 80% of heart diseases, strokes, and type 2 diabetes as well as more than 30% of cancers [1]. These behaviors explain 50% of premature mortality and morbidity in the United States [2]. Improving the efficacy of the interventions dedicated to sustainably change these behaviors will have a greater impact on population health than any therapeutic strategy [3]. One of the possible solutions is the use of nonpharmacological interventions such as digital health interventions [4]. Among eHealth tools, mHealth includes mobile phone health apps and connected health devices [5,6]. Between 20% and 80% of adults are equipped with connected health devices, from one country to another [7-11]. Behavioral intervention technologies (BITs) "employ a broad range of technologies, such as mobile phones, the Web, and sensors, to support users in changing behaviors and cognitions related to health, mental health, and wellness" [12]. Digital health interventions open up new opportunities in the area of prevention and therapeutics: quantified-self activities and behaviors; sharing of clinical, psychological, and behavioral data (ie, short message service [SMS] and social networks); real-time analysis of health data; delivering of health promotion messages (ie, mobile phone and Web apps); involvement of health professionals (ie, telemedicine); e-coaching; social support of the family; and peers in the social environment [13,14]. Several observational studies indicate benefits of BITs on health behavior change in patients with chronic disease or at-risk people using interventions to assist with stronger drug compliance, smoking cessation, alcohol consumption reduction, weight management, and better diets [15-18]. Does the existence of a benefit provide evidence of an effective and safe tool and the only approach to generalize it?

The evaluation of BITs before market and their ongoing monitoring remain questionable, especially in a world propelled by marketing strategies and the lack of regulation of health solutions that do not belong to the category of medical devices. According to the US Food and Drug Administration (FDA) [19,20], the validation process for drugs and implantable medical devices involves the "collection and evaluation of data, from the process design stage through commercial production, which establishes scientific evidence that a process is capable of consistently delivering quality product." To assess the efficacy and safety of a health product, there is a need to use a specific scientific paradigm, which is "a set of principles and methods shared by a scientific community" [21]. A framework is a model for planning processes or for action plans, which brings a systematic approach to developing, managing, and evaluating interventions [22]. Regulators, researchers, and manufacturers share a consensual paradigm for drugs [19,20]. This *clinical trial* framework guides the development from lab to authorization and monitoring. It is organized in five phases: Phase 0 (ie, preclinical) to identify mechanisms; Phase 1 to determine tolerance in healthy humans; Phase 2 to identify the optimal dose for a small number of patients (ie, pilot trial); Phase 3 to demonstrate evidence of efficacy and safety (ie, randomized controlled trial); and Phase 4 to ensure long-term safety [23]. The number of studies assessing BITs is growing exponentially. However, their methodological designs do not follow an established validation framework. They flow from individual or research team choice, context, and/or opportunity.

The objective of this study was to identify the frameworks for the validation and monitoring of BITs proposed in the literature, besides the five-phase drug development model, and to categorize them.

Methods

Data Collection

We have conducted a narrative literature review of articles published up to April 4, 2019, of the validation and monitoring frameworks of BITs. The main databases that we searched were MEDLINE, PsycINFO, Web of Science, Science Direct, and the Journal of Medical Internet Research (JMIR) database. They have been chosen as reference databases in biomedicine and psychology and comprise the largest general scientific database [24]. The search keywords were as follows: ("model" OR "framework" OR "process" OR "evaluation process" OR "validation process") AND ("behavior" OR "behavior change" OR "behavioral intervention") AND ("digital health" OR "ehealth" OR "mhealth" OR "connected health" OR "medical app" OR "smartphone" OR "iphone" OR "email" OR "text message" OR "SMS" OR "mobile app" OR "smartphone app" OR "connected health" OR "wearable"). New articles were extracted based on analyzed articles.

Description of Frameworks

The frameworks were listed in chronological order. Items for each framework were as follows: name of the framework—if none was given, a phrase describing it in the original article was written within quotation marks; publication year; name of the

first author, called the *creator* here; country of the creator; organization having funded the creator; health focus or lack of health focus; the target group (ie, population for whom the framework had been initially designed, for example, researchers, health professionals, or software designers); and its design. For the latter, we noted the involvement of the final users (ie, early stage or systematic) and the chain of the development stages (ie, linear, iterative, evolutive, and/or concurrent) [25]. A linear process executes the different stages in a sequential order. An iterative process combines one or more stages before chaining. An evolutive process executes them in a circular pattern in order to obtain, with each revolution, a more mature version of the product. Finally, a parallel process executes one stage or more in a concurrent way. If a framework had evolved through time and had been described by a new publication, the latter was also written in the synthesis table but its creator, as well as the creator's country and funding organization, remained the original one.

Categorization of Frameworks and Paradigms

The frameworks were categorized according to the accepted procedure of translational research [26,27]. It may be described as a continuum of steps. It starts with a *prototyping* step, which has a background in engineering, followed by a *mechanisms* step based on a theoretical approach to health behavior change. *Concept* follows, which is a proof-of-concept step regarding health impact based on an exploratory intervention trial. This is followed by a demonstration step, *evidence* of health efficacy and effectiveness, which is based on a confirmatory interventional trial. Finally, a *surveillance* step is implemented for market use based on implementation and dissemination studies.

The frameworks follow three general paradigms: (1) the biomedical paradigm, with its essential phase of clinical research to identify the benefits and risks; (2) the engineering paradigm, with its essential phase to improve the device; and (3) the behavioral paradigm, with its essential phase to evaluate the impact on health behaviors.

Results

Overview

The literature review identified 46 frameworks, besides the five-phase drug development model, that met the research criteria (see Table 1).

Table 1. Validation and monitoring frameworks of BITs. BIT: behavioral intervention technology.

Frameworks	Authors, year
1. Waterfall model	Royce, 1970 [28]
2. PRECEDE-PROCEED[a] model	Green, 1974 [29]
3. Prototyping model	Floyd, 1984 [30]
4. Five-phase cancer control model	Greenwald and Cullen, 1985 [31]
5. Flay's eight-stage health promotion model	Flay, 1986 [32]
6. V life cycle model	Rook, 1986 [33]
7. Spiral life cycle model	Boehm, 1988 [34]
8. Star life cycle model	Harston and Dix, 1989 [35]
9. Rapid application development	Martin, 1991 [36]
10. National Institute on Drug Abuse's (NIDA) stage model	Onken et al, 1997 [37]
11. Intervention mapping	Bartholomew et al, 1998 [38]
12. Usability engineering life cycle	Mayhew, 1999 [39]
13. Agile software management	Beck et al, 2001 [40]
14. Information technology (IT) implementation framework	Kukafka et al, 2003 [41]
15. Multiphase Optimization STrategy (MOST)	Collins et al, 2005 [42]
16. Framework for evaluating emergent eHealth resources	Pagliari, 2007 [43]
17. Consolidated Standards of Reporting Trials (CONSORT) statement for nonpharmacologic treatments	Boutron et al, 2008 [44] (updated in Boutron et al, 2017 [45])
18. Iterative and incremental model	Cockburn, 2008 [46]
19. Medical Research Council (MRC) complex intervention	Craig et al, 2008 [47]
20. eHealth interventions evaluation process	Catwell and Sheikh, 2009 [48]
21. Center for eHealth Research (CeHRes) roadmap for the development of eHealth technologies	Van Gemert-Pijnen et al, 2011 [49], and Van Velsen et al, 2013 [50]
22. The behavior change wheel	Michie et al, 2011 [51]
23. Consolidated Standards of Reporting Trials of Electronic and Mobile HEalth Applications and onLine TeleHealth (CONSORT-EHEALTH)	Eysenbach et al, 2011 [52], and Eysenbach et al, 2013 [53]
24. mHealth development and evaluation framework	Whittaker et al, 2012 [17]
25. Explore Values, Operationalize and Learn, and eValuate Efficacy (EVOLVE) mixed-methods model	Peterson et al, 2013 [54]
26. Development process of Young and Active	Riiser et al, 2013 [55]
27. It's LiFe! User-centered design process	Van der Weegen et al, 2013 [56]
28. DoTTI[b] development framework	Smits et al, 2014 [57]
29. National Institutes of Health (NIH) Stage model	Onken et al, 2014 [58]
30. Behavioral intervention technology model	Mohr et al, 2014 [12]
31. Five-step content validity process	Kassam-Adams et al, 2015 [59]
32. Steps for developing a text-messaging program	Abroms et al, 2015 [60]
33. Person-based approach	Riiser et al, 2013 [55], and Yardley et al, 2015 [61]
34. Obesity-Related Behavioral Intervention Trials (ORBIT) model	Van der Weegen et al, 2013 [56], and Czajkowski et al, 2015 [62]
35. Pragmatic Framework for developing just-in-time adaptive interventions (JITAIs)	Smits et al, 2014 [57], and Nahum-Shani et al, 2015 [63]
36. TElehealth in CHronic disease (TECH) conceptual model	Salisbury et al, 2015 [64]
37. Network for the Improvement of Addiction Treatment (NIATx) model	Gustafson et al, 2016 [65]
38. Integrate, Design, Assess, and Share (IDEAS) framework	Mummah et al, 2016 [66]

Frameworks	Authors, year
39. Chronic disease mHealth app intervention design framework	Wilhide III et al, 2016 [67]
40. Three-phase human-centered design methodology	Harte et al, 2017 [68]
41. DREAM-GLOBAL[c] framework	Maar et al, 2017 [69]
42. Processes and recommendations for creating mHealth apps for low-income populations	Stephan et al, 2017 [70]
43. Accelerated Creation-To-Sustainment (ACTS) model	Mohr et al, 2017 [71]
44. User-centered design process	Vilardaga et al, 2018 [72]
45. Eight-step scoping framework	Davidson et al, 2019 [73]
46. Targeting, Understanding, Designing, Evaluating, and Refining (TUDER) framework	Wang et al, 2019 [74]

[a]PRECEDE-PROCEED: Predisposing, Reinforcing, and Enabling Constructs in Educational Diagnosis and Evaluation—Policy, Regulatory, and Organizational Constructs in Educational and Environmental Development.

[b]DoTTI: Design and develOpment, Testing early iterations, Testing for effectiveness, Integration and implementation.

[c]DREAM-GLOBAL: Diagnosing hypeRtension—Engaging Action and Management in Getting LOwer Bp in Aboriginal and LMIC (lower- and middle-income countries).

Description of Frameworks

Frameworks

The results showed that, out of 46 frameworks, 2 (4%) were designed in the 1970s, 6 (13%) in the 1980s, 4 (9%) in the 1990s, 8 (17%) in the 2000s, and 26 (57%) in the 2010s (see Multimedia Appendix 1). Their creators were mainly from the United States (24/46, 52%) and the United Kingdom (10/46, 21%). The Netherlands, Germany, Canada, Australia, Brazil, France, Ireland, Norway, and New Zealand were the remaining creators' countries. A total of 8 frameworks out of 46 (17%) were created by private software company workers. A total of 4 frameworks out of 46 (9%) were designed by a national or a supranational public organization: the National Institutes of Health (NIH), the Medical Research Council (MRC), the National Health Service (NHS), and the Consolidated Standards of Reporting Trials (CONSORT). Universities were the main institutions (34/46, 74%), with 16 out of 46 (35%) from the United States. A total of 29 out of 46 frameworks (63%) were supported by a research grant: out of 46 frameworks, 29 (63%) were funded by a public grant, 12 (26%) were funded by the NIH, and 6 (13%) were funded by the English public health system. Out of 46 frameworks, 5 (11%) were funded by universities, 3 (7%) were funded by the private sector, and 3 (7%) were funded by a public-private partnership.

Purpose and Target Population

A total of 36 frameworks out of 46 (78%) were created with an individual health focus. Among them, the NIH Stage model was created to be relevant for clinical sciences [58] and the CONSORT statement was created for nonpharmacologic treatments for all nonpharmacological interventions [44]. Out of 46 frameworks, 3 (7%) aimed to validate health promotion programs [29,32,38]. A total of 23 of 46 (50%) were targeted to validate health behavior change interventions: of these 23, 15 (65%) were aimed at eHealth interventions in general [12,41,43,48,49,52,59,61,64-66,71-74], 2 (9%) at Internet interventions [55,57], and 6 (26%) at mobile interventions [17,56,60,67,69,70]. Among the 8 frameworks out of 23 (35%) aimed at health behavior change that were not focused on eHealth, 5 (63%) were created for behavior change in general [42,47,51,54,63] and 3 (38%) were dedicated to diseases such as cancer [31], addiction [37], and all chronic disease [62]. A total of 10 out of 46 frameworks (22%) were created with an original purpose that was not associated with health, but was associated with software development [28,33,34,36,40,46], human-computer interfaces [35,39], or engineering in general [30,68].

Out of 46 frameworks, 17 (37%) were designed explicitly for intervention designers: 11 (24%) for software designers [17,30,33,35,36,39,43,46,48,49,64], 1 of which was for software evaluators [66], and 6 (13%) for designers of health behavior change interventions [12,41,51,61-63], 2 of which were based on technologies to change health behaviors [12,41]. A total of 9 out of 46 frameworks (20%) were aimed at researchers; these included researchers in general [54,58], scientific paper authors or scientific journal editors [44,51], researchers and stakeholders [32,71], researchers and industry professionals [66], and health researchers and software designers [43,60]. Finally, 13 frameworks out of 46 (28%) were intended for all stakeholders: professionals, researchers, users, clinicians, and other health helpers [42,46,55-57,59,65,67,68-70,72,73]. In 6 out of 46 cases (13%), the target of the framework was not mentioned; however, the reading of these articles was directed toward software designers [28,34,40], researchers, stakeholders, and health program planners [29,31,37]. Out of 46, 25 frameworks (54%) were created for software designers and 24 (52%) were created for researchers.

Organization

In 28 out of 46 frameworks (61%), the final user was involved early and systematically (ie, at each step). In 27 of them (59%), he or she played an active role in the BIT's specification and assessment [17,35,36,39,40,43,48,49,54-57,59-62,64-74]. In the V life cycle model, the role was only as an evaluator [33].

A total of 2 frameworks out of 46 (4%) had a linear sequence in their phases. The Waterfall model adopts a seven-phase

structure: identification of the system specifications, identification of the software specifications, analysis, program design, coding, tests, and operations [28]. Inspired by the latter, the V life cycle model has a V-shaped structure, which matches development and testing phases and involves clients, followed by final users, in the development [33]. In addition to the linear sequence, 17 out of 46 frameworks (37%) had an iterative structure [12,17,31,32,42,44,52,54-56,61,62,66,69,70,72,73], allowing a refining step at each phase in the case of suboptimal results. The iterative structure is one of the assets used by the Multiphase Optimization STrategy (MOST) and the Obesity-Related Behavioral Intervention Trials (ORBIT) model frameworks to optimize interventions [42,62]. A total of 15 out of 46 frameworks (33%) were also associated with an evolutive process, creating a cyclic organization to improve the product progressively [30,34,35,37-39,43,47-49,51,58,60,67,68]. Finally, 11 out of 46 frameworks (24%) also integrated a parallel process [29,36,40,41,46,57,59,64,65,71,74]; 1 of those was the Agile software management framework, which aims for a high degree of adaptability to satisfy the clients [40].

Framework Categorization

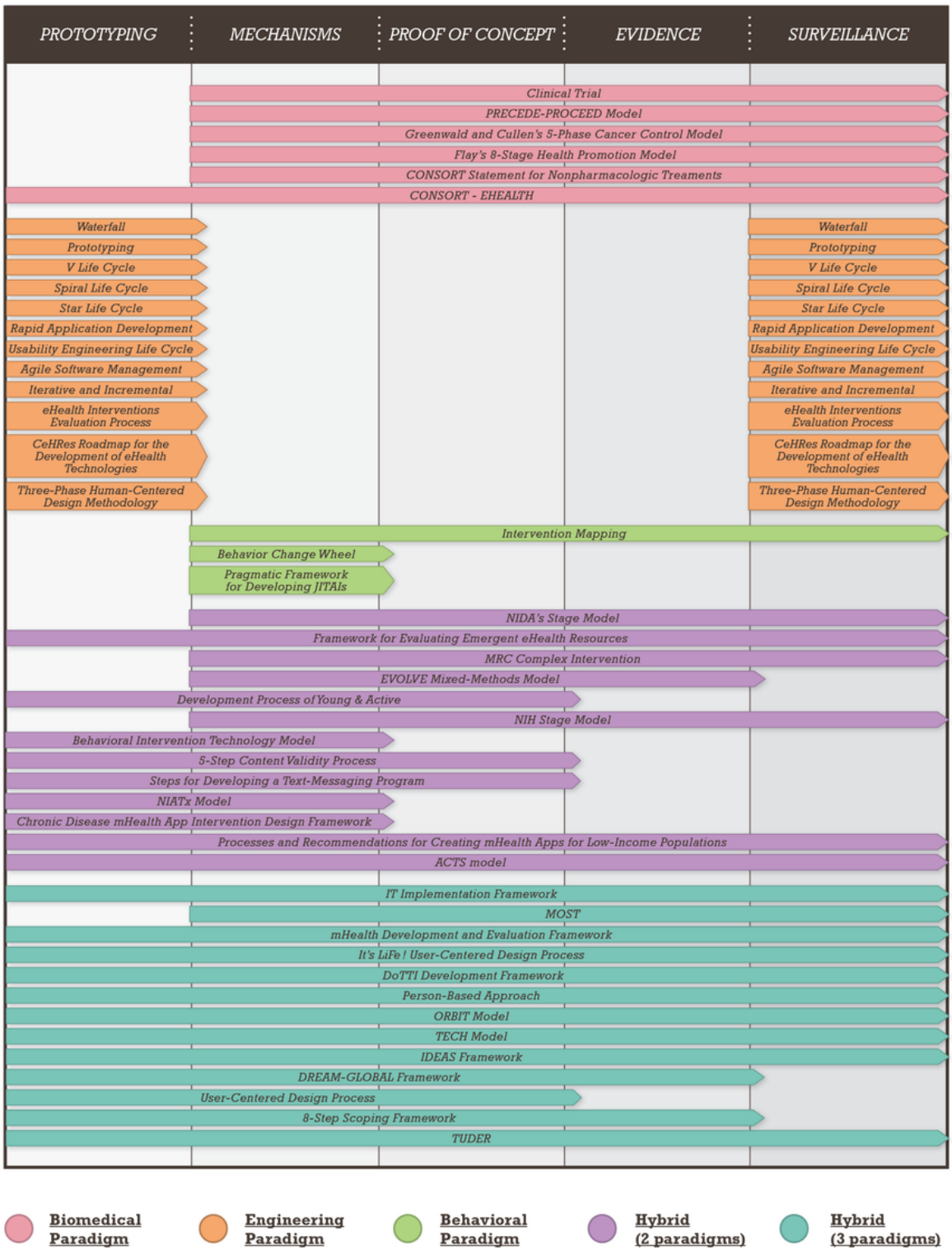
The categorization of frameworks, according to their coverage of the continuum of steps, is shown in Figure 1.

The Waterfall model is the only one that covers only the prototyping step [28]. Other frameworks cover the prototyping and surveillance steps by integrating the final user [30,33-36,39,40,46,48,49,68].

The behavior change wheel and the pragmatic framework for developing just-in-time adaptive interventions (JITAIs) are focused on the mechanisms step [51,63]. The BIT model, the chronic disease mHealth app intervention design framework, and the Network for the Improvement of Addiction Treatment (NIATx) model cover the prototyping and mechanisms steps by associating the conception of both numeric and behavioral interventions [12,17,65].

The Explore Values, Operationalize and Learn, and eValuate Efficacy (EVOLVE) and MOST frameworks cover the mechanisms, concept, and evidence steps. A total of 9 frameworks out of 46 (20%) are extended over the mechanisms, concept, evidence, and surveillance steps [29,31,32,37,44,47,49,58,62]. Finally, the whole continuum of steps is covered by 12 out of 46 (26%) frameworks [17,41,43,52,56,57,61,64,66,70,71,74].

Figure 1. Proposed frameworks to validate and monitor behavioral intervention technologies (BITs). ACTS: Accelerated Creation-To-Sustainment; CeHRes: Center for eHealth Research; DoTTI: Design and develOpment, Testing early iterations, Testing for effectiveness, Integration and implementation; DREAM-GLOBAL: Diagnosing hypeRtension—Engaging Action and Management in Getting LOwer Bp in Aboriginal and LMIC (lower- and middle-income countries); EVOLVE: Explore Values, Operationalize and Learn, and eValuate Efficacy; IDEAS: Integrate, Design, Assess, and Share; IT: information technology; JITAI: just-in-time adaptive intervention; MOST: Multiphase Optimization STrategy; MRC: Medical Research Council; NIATx: Network for the Improvement of Addiction Treatment; NIDA: National Institute on Drug Abuse; NIH: National Institutes of Health; ORBIT: Obesity-Related Behavioral Intervention Trials; PRECEDE-PROCEED: Predisposing, Reinforcing, and Enabling Constructs in Educational Diagnosis and Evaluation—Policy, Regulatory, and Organizational Constructs in Educational and Environmental Development; TECH: TElehealth in CHronic disease; TUDER: Targeting, Understanding, Designing, Evaluating, and Refining.



Paradigm Categorization

Figure 1 illustrates the categorization of frameworks based on their paradigms. Most belonged to one paradigm (ie, biomedical, engineering, or behavioral), while few referred to two or three paradigms.

Out of 46 frameworks, 12 (26%) fell within the paradigm of engineering [28,30,33-36,39,40,46,48,49,68] and 3 (7%) were based on the behavioral paradigm [38,51,63]. A total of 9 (20%) frameworks were based on these two latter paradigms [12,55,59,60,65,67,70-72]. A total of 5 (11%) frameworks were based on the biomedical paradigm, which draws on the clinical trial, specifically the randomized controlled trial [28,31,32,44,52]. A total of 4 others (9%) also incorporated a theoretical approach of behavior change [37,47,54,58]. The framework for evaluating emergent eHealth resources was based on the biomedical and engineering paradigms [43]. A total of 12 (26%) frameworks mixed three paradigms [17,41,42,56,57,61-63,67,69,73,74].

Discussion

Principal Findings

Our narrative review of the literature showed the absence of a unique and consensual validation and surveillance framework for health behavior intervention technology. This conclusion is in keeping with the statement of Bradway et al about mHealth assessment: “too many initiatives, too few answers” [75]. To date, 46 frameworks coexist, besides the five-phase drug development model, without a predominance of, or a convergence toward, one of them. Their number has increased exponentially over time since the 1970s. The United States and the United Kingdom are the main countries that have proposed models, probably due to their university research productivity and new-technology industry leadership [76]. They are motivated, on the one hand, by the plethoric and easily available offer of digital tools and networks and, on the other hand, by the failure to prevent unhealthy diet, addiction, and physical inactivity [2].

Beyond the quantitative increase of these frameworks, this review underlines an increased heterogeneity. One surprising finding is that the benefit-to-risk ratio is not a main goal. Contrary to drug claims regarding societal recognition and reimbursement for their impact on health outcomes [77], showing evidence of safety and efficacy is paradoxically not a priority for BITs. This fact can explain why researchers have noted that medical mobile phone apps are lightly used by patients in the medium term, have little involvement from health professionals, lack characterization of their content, and have a low integration of behavior change theories [78]. The same problem can be found for connected health devices. The lack of evidence of their efficacy leads to their nonprescription by general practitioners, who picture them mainly as a supplementary organizational constraint [79]. Most frameworks are inspired by the paradigms of engineering and behavioral sciences. Engineering is focused on technicity (ie, release criteria, ergonomics, user experience, and data security), which leads to uncertainties about these technologies’ health contributions and turns them into entertainment products or gadgets [80]. The behavioral sciences bring an understanding of behavior change mechanisms and skills techniques without determining and comparing their health impacts [81,82]. This explains how translational research is unusual in this area compared to that for drugs [83]. To date, for instance, there is no validation model for psychotherapies [84].

Beyond the quantitative increase and diversification of frameworks, we have noted the emergence of common principles. The first one is an assumed superiority with respect to health and health behavior change of some BITs compared to others. This hypothesis aims to address the plethoric and exponential offer of digital solutions sustained by marketing. This growing need to know the efficacy of each BIT and to compare them to other solutions to change health behaviors comes from both health professionals, who seek to recommend or even prescribe evidence-based solutions, and health users, who seek solutions beyond marketing and/or users’ experiences from social networks.

The second is the requirement of user involvement during the beginning of BIT development and after market access. This principle exists in every recent framework based on the behavioral science, engineering, or biomedical paradigms. It requires a coconstruction with the user and considers his or her experience. It involves the end user through a user-centered design from engineering [85,86]. It requires a patient-centered approach to the provision of personalized care in the biomedical area [87]. The patient’s preferences are considered and may allow improved chances of stable and efficient use.

The third is the ambition to evaluate BITs beyond simple satisfaction. This is more or less oriented by manufacturers to aim at effectiveness in medicine becoming more and more predictive, preventive, personalized, and participative; this is called 4P medicine [88,89].

The fourth is the shortening of the delay for upstream validation, until its complete suppression in some engineering frameworks, and a multiplication of the downstream surveillance methods, such as implementation studies, experience feedback (eg, to a learned society or health authority), and big data analyses combining multimodal data. The shortening of the upstream period is linked, on the one hand, to a supposedly low dangerousness of these health solutions and, on the other hand, to the short life cycle of the technologies involved [90]. The development time of a drug, from the lab bench to its marketing authorization, is between 17 and 24 years [91-93]. It is incompatible with digital innovation. Digital industrialists assume that hundreds of millions of users will allow for the evaluation of BITs’ effectiveness on a large scale thanks to big data analyses [94,95].

The fifth is the introduction of hybrid frameworks integrating the development and updating processes from engineering, behavior change theories, techniques from behavioral sciences, and the rigorous approach of validation from the biomedical area (eg, the FDA and the Environmental Management Association). This hybridization contributes to the creation of interventions standing between medical devices and products in the category of *general commodities*. The scientific process to develop and validate them, the rules for marketing

authorization, and surveillance have not yet been defined for a country or a continent and are still under construction. Many digital industrialists wish to avoid the validation process for medical devices in view of its constraints and costs; at the same time, they aim to indicate to health professionals and users the health impact of their solutions. They are encouraged to address this by mutual insurances, other insurances, and banks.

We stress two extreme positions for the studied frameworks that may become problematic for the development of BITs in the health sector. The first is that of permanent technological innovation. It is embodied in the Agile software management framework. Flexibility, based on user demand, dictates each evolution. Health is a market like any other. If this approach was to win in this area, for example, to make substantial savings in clinical research, no counterpart could be asked for, concerning health benefits. Big data analyses will not guarantee a relevant comparability between BITs [96]. The second problem at the opposite extreme is that marketing authorization of the BIT must be conditional on the completion of a well-designed randomized controlled trial. This process has been a part of the success of drugs in the last century. These trials are long, costly, and debatable in terms of methodology for BITs, if only for the choice of control group [97].

Limitations and Strengths

To identify every validation and surveillance framework of BITs is challenging. This area is new, it is based on a number of different paradigms, and develops in an uncontrolled way from one country to another and from one industry to another. One of the identified difficulties within these publications is the lack of common terminology [98,99]. The originality of our review is the spanning of the three approaches (ie, engineering, behavioral sciences, biomedical sciences). Although they are often opposed, the concern here was to note the borrowings and contributions that they could mutually bring through transversal thinking.

Conclusions

Validating a BIT for health has the following specific challenges: achieve an efficient and quick development; understand and promote long-term adherence; be supported by behavior change theories and techniques; evaluate effectiveness and cost-effectiveness; and ensure rigorous management in terms of regulation, ethics, and information [93]. Patient expectations are relayed through the media and through the arrival on the health market of digital players proposing short life-cycle products due to rapid obsolescence of technology; this justifies the need for a consensual validation and surveillance framework for nonpharmacological interventions. This framework will necessarily differ from that used for drugs. Our review has identified 46 frameworks, none of which dominates for BITs. Three paths are opened: (1) the anarchic continuous development of new competing frameworks that prevents any convergence in a standardized validation and surveillance framework, with its consequential recognition by health authorities; (2) the movement toward the type of framework for drug evaluation centered on establishing its effectiveness before marketing authorization to guarantee its safety for users, which is heavy and costly; and (3) the implementation of a framework reliant on big data analysis, based on a posteriori research and an autoregulation of a market; however, that does not address the safety risk for the health user, as the market will not regulate safety or efficacy issues. This article calls for a minimal upstream clinical phase and an increased surveillance of BITs to address the risk of semantic amalgams, inappropriate prescriptions, and induced misuses [16,100]. A BIT cannot be a simple tool but must be a complex strategy integrated in a given environment [101]. A BIT must make sense by virtue of its interaction with the context, building a system, and must be evaluated as such [100].

Acknowledgments

The authors would like to thank the members of the Plateforme universitaire Collaborative d'Evaluation des programmes de Prévention et de Soins de support (CEPS) for their help in better defining and using nonpharmacological interventions for patients and the SIRIC Montpellier Cancer (Grant INCa Inserm DGOS 12553).

Conflicts of Interest

None declared.

Multimedia Appendix 1
Table of frameworks.
[PDF File (Adobe PDF File), 225 KB - jmir_v21i10e13606_app1.pdf]

References

1. Global Action Plan for the Prevention and Control of Noncommunicable Diseases 2013-2020. Geneva, Switzerland: World Health Organization; 2013. URL: http://apps.who.int/iris/bitstream/handle/10665/94384/9789241506236_eng.pdf?sequence=1. Archived [accessed 2019-09-06] [WebCite Cache ID 723P8FQzi]
2. Institute of Medicine, Committee on Quality of Health Care in America. Crossing The Quality Chasm: A New Health System for the 21st Century. Washington, DC: National Academies Press; 2001.

3. Sabaté E, De Geest S. Adherence to long-term therapies management: A call for cardiovascular nursing managers and policymakers. Prog Cardiovasc Nurs 2004;19(1):28-29. [Medline: 15017153]
4. CEPS Platform. URL: https://www.iceps2020.com/iceps-congres/ceps-platform/ [accessed 2019-09-12] [WebCite Cache ID 6rjFd3LwA]
5. Fifty-Eighth World Health Assembly. Geneva, Switzerland: World Health Organization; 2005. URL: http://apps.who.int/gb/ebwha/pdf_files/WHA58-REC1/english/A58_2005_REC1-en.pdf [accessed 2019-09-06]
6. World Health Organization. mHealth: New Horizons for Health Through Mobile Technologies. Global Observatory for eHealth Series - Volume 3. Geneva, Switzerland: World Health Organization; 2011.
7. Gandhi M, Wang T. Rock Health. 2015. Digital health consumer adoption: 2015 URL: https://rockhealth.com/reports/digital-health-consumer-adoption-2015/ [accessed 2019-09-06] [WebCite Cache ID 6rjNqYuuL]
8. Fox S, Duggan M. Pew Research Center. Washington, DC: Pew Internet & American Life Project; 2012 Nov 08. Mobile health 2012 URL: http://www.pewinternet.org/2012/11/08/mobile-health-2012 [accessed 2019-09-06] [WebCite Cache ID 6rjO2BIjg]
9. Kontos E, Blake KD, Chou WS, Prestin A. Predictors of eHealth usage: Insights on the digital divide from the Health Information National Trends Survey 2012. J Med Internet Res 2014 Jul 16;16(7):e172 [FREE Full text] [doi: 10.2196/jmir.3117] [Medline: 25048379]
10. mHealth App Developer Economics 2016: The Current Status and Trends of the mHealth App Market. Berlin, Germany: Research 2 Guidance; 2016 Oct. URL: http://research2guidance.com/wp-content/uploads/2016/10/mHealth-App-Developer-Economics-2016-v17-Preview-1-1.pdf [accessed 2019-09-06] [WebCite Cache ID 6rsASIhiB]
11. Danova T. Business Insider France. 2013 Oct 02. Morgan Stanley: 75 billion devices will be connected to the Internet of Things by 2020 URL: http://www.businessinsider.fr/us/75-billion-devices-will-be-connected-to-the-internet-by-2020-2013-10/ [accessed 2019-09-06] [WebCite Cache ID 6udcbftF4]
12. Mohr DC, Schueller SM, Montague E, Burns MN, Rashidi P. The behavioral intervention technology model: An integrated conceptual and technological framework for eHealth and mHealth interventions. J Med Internet Res 2014 Jun 05;16(6):e146 [FREE Full text] [doi: 10.2196/jmir.3077] [Medline: 24905070]
13. Klasnja P, Pratt W. Healthcare in the pocket: Mapping the space of mobile-phone health interventions. J Biomed Inform 2012 Feb;45(1):184-198 [FREE Full text] [doi: 10.1016/j.jbi.2011.08.017] [Medline: 21925288]
14. Petit A, Cambon L. Exploratory study of the implications of research on the use of smart connected devices for prevention: A scoping review. BMC Public Health 2016 Jul 11;16:552 [FREE Full text] [doi: 10.1186/s12889-016-3225-4] [Medline: 27401769]
15. Free C, Phillips G, Galli L, Watson L, Felix L, Edwards P, et al. The effectiveness of mobile-health technology-based health behaviour change or disease management interventions for health care consumers: A systematic review. PLoS Med 2013;10(1):e1001362 [FREE Full text] [doi: 10.1371/journal.pmed.1001362] [Medline: 23349621]
16. Perski O, Blandford A, West R, Michie S. Conceptualising engagement with digital behaviour change interventions: A systematic review using principles from critical interpretive synthesis. Transl Behav Med 2017 Jun;7(2):254-267 [FREE Full text] [doi: 10.1007/s13142-016-0453-1] [Medline: 27966189]
17. Whittaker R, Merry S, Dorey E, Maddison R. A development and evaluation process for mHealth interventions: Examples from New Zealand. J Health Commun 2012;17 Suppl 1:11-21. [doi: 10.1080/10810730.2011.649103] [Medline: 22548594]
18. Lau PW, Lau EY, Wong DP, Ransdell L. A systematic review of information and communication technology-based interventions for promoting physical activity behavior change in children and adolescents. J Med Internet Res 2011 Jul 13;13(3):e48 [FREE Full text] [doi: 10.2196/jmir.1533] [Medline: 21749967]
19. Agalloco J. Validation: An unconventional review and reinvention. PDA J Pharm Sci Technol 1995;49(4):175-179. [Medline: 7552236]
20. Process Validation: General Principles and Practices. Silver Spring, MD: US Food and Drug Administration; 2011 Jan. URL: https://www.fda.gov/media/71021/download [accessed 2019-09-06] [WebCite Cache ID 6smyYNQ2g]
21. Kuhn TS. The Structure of Scientific Revolutions. Chicago, IL: University Of Chicago Press; 1996.
22. Green L, Kreuter M. Health Program Planning: An Educational and Ecological Approach. 4th edition. New York, NY: McGraw-Hill Humanities/Social Sciences/Languages; 2004.
23. ClinicalTrials.gov. Phase URL: https://clinicaltrials.gov/ct2/help/glossary/phase [accessed 2019-09-06] [WebCite Cache ID 6sn4cZsIA]
24. Chapman D. Health-related databases. J Can Acad Child Adolesc Psychiatry 2009 May;18(2):148-149 [FREE Full text] [Medline: 19495437]
25. Pressman RS. Software Engineering: A Practitioner's Approach. 5th edition. New York, NY: McGraw-Hill Education; 2001.
26. Fang FC, Casadevall A. Lost in translation: Basic science in the era of translational research. Infect Immun 2010 Feb;78(2):563-566 [FREE Full text] [doi: 10.1128/IAI.01318-09] [Medline: 20038540]
27. Fishbein DH, Ridenour TA, Stahl M, Sussman S. The full translational spectrum of prevention science: Facilitating the transfer of knowledge to practices and policies that prevent behavioral health problems. Transl Behav Med 2016 Mar;6(1):5-16 [FREE Full text] [doi: 10.1007/s13142-015-0376-2] [Medline: 27012249]

28. Royce WW. Managing the development of large software systems: Concepts and techniques. In: Proceedings of the 9th International Conference on Software Engineering. Los Alamitos, CA: IEEE Computer Society Press; 1987 Presented at: 9th International Conference on Software Engineering; March 30-April 2, 1987; Monterey, CA p. 328-338 URL: http://www-scf.usc.edu/~csci201/lectures/Lecture11/royce1970.pdf
29. Green L. Toward cost-benefit evaluations of health education: Some concepts, methods, and examples. Health Educ Monogr 2016 Aug 27;2(1_suppl):34-64 [FREE Full text] [doi: 10.1177/10901981740020S106]
30. Floyd C. A systematic look at prototyping. In: Budde R, Kuhlenkamp K, Mathiassen L, Züllighoven H, editors. Approaches to Prototyping. Berlin, Germany: Springer-Verlag Berlin Heidelberg; 1984:1-18.
31. Greenwald P, Cullen J. The new emphasis in cancer control. J Natl Cancer Inst 1985 Mar;74(3):543-551. [Medline: 3883037]
32. Flay BR. Efficacy and effectiveness trials (and other phases of research) in the development of health promotion programs. Prev Med 1986 Sep;15(5):451-474. [Medline: 3534875]
33. Rook P. Controlling software projects. Softw Eng J 1986 Jan;1(1):7-16. [doi: 10.1049/sej.1986.0003]
34. Boehm BW. A spiral model of software development and enhancement. Comput 1988 May;21(5):61-72. [doi: 10.1109/2.59]
35. Hartson H, Hix D. Toward empirically derived methodologies and tools for human-computer interface development. Int J Man Mach Stud 1989 Oct;31(4):477-494. [doi: 10.1016/0020-7373(89)90005-9]
36. Martin J. Rapid Application Development. Indianapolis, IN: Macmillan Coll Div; 1991.
37. Onken LS, Blaine JD, Battjes RJ. Behavioral therapy research: A conceptualization of a process. In: Henderson SW, Santos AB, editors. Innovative Approaches for Difficult-to-Treat Populations. Washington, DC: American Psychiatric Press; 1997:477-485.
38. Bartholomew LK, Parcel GS, Kok G. Intervention mapping: A process for developing theory- and evidence-based health education programs. Health Educ Behav 1998 Oct;25(5):545-563. [doi: 10.1177/109019819802500502] [Medline: 9768376]
39. Mayhew DJ. The Usability Engineering Lifecycle: A Practitioner's Handbook for User Interface Design (interactive Technologies). London, UK: Morgan Kaufmann Publishers; 1999.
40. Beck K, Beedle M, Bennekum A, Cockburn A, Cunningham W, Fowler M, et al. agilemanifesto.org. 2001. Manifesto for Agile software development URL: http://agilemanifesto.org/ [accessed 2019-09-06] [WebCite Cache ID 723NqYRGK]
41. Kukafka R, Johnson SB, Linfante A, Allegrante JP. Grounding a new information technology implementation framework in behavioral science: A systematic analysis of the literature on IT use. J Biomed Inform 2003 Jun;36(3):218-227 [FREE Full text] [Medline: 14615230]
42. Collins LM, Murphy SA, Nair VN, Strecher VJ. A strategy for optimizing and evaluating behavioral interventions. Ann Behav Med 2005 Aug;30(1):65-73. [doi: 10.1207/s15324796abm3001_8] [Medline: 16097907]
43. Pagliari C. Design and evaluation in eHealth: Challenges and implications for an interdisciplinary field. J Med Internet Res 2007 May 27;9(2):e15 [FREE Full text] [doi: 10.2196/jmir.9.2.e15] [Medline: 17537718]
44. Boutron I, Moher D, Altman DG, Schulz KF, Ravaud P, CONSORT Group. Methods and processes of the CONSORT Group: Example of an extension for trials assessing nonpharmacologic treatments. Ann Intern Med 2008 Feb 19;148(4):W60-W66. [doi: 10.7326/0003-4819-148-4-200802190-00008-w1] [Medline: 18283201]
45. Boutron I, Altman D, Moher D, Schulz K, Ravaud P, CONSORT NPT Group. CONSORT statement for randomized trials of nonpharmacologic treatments: A 2017 update and a CONSORT extension for nonpharmacologic trial abstracts. Ann Intern Med 2017 Jul 04;167(1):40-47. [doi: 10.7326/M17-0046] [Medline: 28630973]
46. Cockburn A. Rochester Institute of Technology. 2008. Using both incremental and iterative development URL: http://www.se.rit.edu/~swen-256/resources/UsingBothIncrementalandIterativeDevelopment-AlistairCockburn.pdf [accessed 2019-09-06] [WebCite Cache ID 723RUNAJQ]
47. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M. Developing and evaluating complex interventions: The new Medical Research Council guidance. Int J Nurs Stud 2013 May;50(5):587-592. [doi: 10.1016/j.ijnurstu.2012.09.010] [Medline: 23159157]
48. Catwell L, Sheikh A. Evaluating eHealth interventions: The need for continuous systemic evaluation. PLoS Med 2009 Aug;6(8):e1000126 [FREE Full text] [doi: 10.1371/journal.pmed.1000126] [Medline: 19688038]
49. van Gemert-Pijnen JE, Nijland N, van Limburg M, Ossebaard HC, Kelders SM, Eysenbach G, et al. A holistic framework to improve the uptake and impact of eHealth technologies. J Med Internet Res 2011 Dec 05;13(4):e111 [FREE Full text] [doi: 10.2196/jmir.1672] [Medline: 22155738]
50. Van Velsen L, Wentzel J, Van Gemert-Pijnen JE. Designing eHealth that matters via a multidisciplinary requirements development approach. JMIR Res Protoc 2013 Jun 24;2(1):e21 [FREE Full text] [doi: 10.2196/resprot.2547] [Medline: 23796508]
51. Michie S, van Stralen MM, West R. The behaviour change wheel: A new method for characterising and designing behaviour change interventions. Implement Sci 2011 Apr 23;6:42 [FREE Full text] [doi: 10.1186/1748-5908-6-42] [Medline: 21513547]
52. Eysenbach G, CONSORT-EHEALTH Group. CONSORT-EHEALTH: Improving and standardizing evaluation reports of Web-based and mobile health interventions. J Med Internet Res 2011 Dec 31;13(4):e126 [FREE Full text] [doi: 10.2196/jmir.1923] [Medline: 22209829]
53. Eysenbach G. CONSORT-EHEALTH: Implementation of a checklist for authors and editors to improve reporting of Web-based and mobile randomized controlled trials. Stud Health Technol Inform 2013;192:657-661. [Medline: 23920638]

54. Peterson JC, Czajkowski S, Charlson ME, Link AR, Wells MT, Isen AM, et al. Translating basic behavioral and social science research to clinical application: The EVOLVE mixed-methods approach. J Consult Clin Psychol 2013 Apr;81(2):217-230 [FREE Full text] [doi: 10.1037/a0029909] [Medline: 22963594]
55. Riiser K, Løndal K, Ommundsen Y, Sundar T, Helseth S. Development and usability testing of an Internet intervention to increase physical activity in overweight adolescents. JMIR Res Protoc 2013 Jan 28;2(1):e7 [FREE Full text] [doi: 10.2196/resprot.2410] [Medline: 23612506]
56. van der Weegen S, Verwey R, Spreeuwenberg M, Tange H, van der Weijden T, de Witte L. The development of a mobile monitoring and feedback tool to stimulate physical activity of people with a chronic disease in primary care: A user-centered design. JMIR Mhealth Uhealth 2013 Jul 02;1(2):e8 [FREE Full text] [doi: 10.2196/mhealth.2526] [Medline: 25099556]
57. Smits R, Bryant J, Sanson-Fisher R, Tzelepis F, Henskens F, Paul C, et al. Tailored and integrated Web-based tools for improving psychosocial outcomes of cancer patients: The DoTTI development framework. J Med Internet Res 2014 Mar 14;16(3):e76 [FREE Full text] [doi: 10.2196/jmir.2849] [Medline: 24641991]
58. Onken LS, Carroll KM, Shoham V, Cuthbert BN, Riddle M. Reenvisioning clinical science: Unifying the discipline to improve the public health. Clin Psychol Sci 2014 Jan 01;2(1):22-34 [FREE Full text] [doi: 10.1177/2167702613497932] [Medline: 25821658]
59. Kassam-Adams N, Marsac M, Kohser K, Kenardy J, March S, Winston F. A new method for assessing content validity in model-based creation and iteration of eHealth interventions. J Med Internet Res 2015 Apr 15;17(4):e95 [FREE Full text] [doi: 10.2196/jmir.3811] [Medline: 25881584]
60. Abroms L, Whittaker R, Free C, Mendel Van Alstyne J, Schindler-Ruwisch J. Developing and pretesting a text messaging program for health behavior change: Recommended steps. JMIR Mhealth Uhealth 2015 Dec 21;3(4):e107 [FREE Full text] [doi: 10.2196/mhealth.4917] [Medline: 26690917]
61. Yardley L, Morrison L, Bradbury K, Muller I. The person-based approach to intervention development: Application to digital health-related behavior change interventions. J Med Internet Res 2015 Jan 30;17(1):e30 [FREE Full text] [doi: 10.2196/jmir.4055] [Medline: 25639757]
62. Czajkowski SM, Powell LH, Adler N, Naar-King S, Reynolds KD, Hunter CM, et al. From ideas to efficacy: The ORBIT model for developing behavioral treatments for chronic diseases. Health Psychol 2015 Oct;34(10):971-982 [FREE Full text] [doi: 10.1037/hea0000161] [Medline: 25642841]
63. Nahum-Shani I, Hekler EB, Spruijt-Metz D. Building health behavior models to guide the development of just-in-time adaptive interventions: A pragmatic framework. Health Psychol 2015 Dec;34S:1209-1219 [FREE Full text] [doi: 10.1037/hea0000306] [Medline: 26651462]
64. Salisbury C, Thomas C, O'Cathain A, Rogers A, Pope C, Yardley L, et al. TElehealth in CHronic disease: Mixed-methods study to develop the TECH conceptual model for intervention design and evaluation. BMJ Open 2015 Feb 06;5(2):e006448 [FREE Full text] [doi: 10.1136/bmjopen-2014-006448] [Medline: 25659890]
65. Gustafson DHJ, Maus A, Judkins J, Dinauer S, Isham A, Johnson R, et al. Using the NIATx model to implement user-centered design of technology for older adults. JMIR Hum Factors 2016 Jan 14;3(1):e2 [FREE Full text] [doi: 10.2196/humanfactors.4853] [Medline: 27025985]
66. Mummah SA, Robinson TN, King AC, Gardner CD, Sutton S. IDEAS (Integrate, Design, Assess, and Share): A framework and toolkit of strategies for the development of more effective digital interventions to change health behavior. J Med Internet Res 2016 Dec 16;18(12):e317 [FREE Full text] [doi: 10.2196/jmir.5927] [Medline: 27986647]
67. Wilhide III CC, Peeples MM, Anthony Kouyaté RC. Evidence-based mHealth chronic disease mobile app intervention design: Development of a framework. JMIR Res Protoc 2016 Feb 16;5(1):e25 [FREE Full text] [doi: 10.2196/resprot.4838] [Medline: 26883135]
68. Harte R, Glynn L, Rodríguez-Molinero A, Baker P, Scharf T, Quinlan L, et al. A human-centered design methodology to enhance the usability, human factors, and user experience of connected health systems: A three-phase methodology. JMIR Hum Factors 2017 Mar 16;4(1):e8 [FREE Full text] [doi: 10.2196/humanfactors.5443] [Medline: 28302594]
69. Maar M, Yeates K, Perkins N, Boesch L, Hua-Stewart D, Liu P, et al. A framework for the study of complex mHealth interventions in diverse cultural settings. JMIR Mhealth Uhealth 2017 Apr 20;5(4):e47 [FREE Full text] [doi: 10.2196/mhealth.7044] [Medline: 28428165]
70. Stephan L, Dytz Almeida E, Guimaraes R, Ley A, Mathias R, Assis M, et al. Processes and recommendations for creating mHealth apps for low-income populations. JMIR Mhealth Uhealth 2017 Apr 03;5(4):e41 [FREE Full text] [doi: 10.2196/mhealth.6510] [Medline: 28373155]
71. Mohr D, Lyon A, Lattie E, Reddy M, Schueller S. Accelerating digital mental health research from early design and creation to successful implementation and sustainment. J Med Internet Res 2017 May 10;19(5):e153 [FREE Full text] [doi: 10.2196/jmir.7725] [Medline: 28490417]
72. Vilardaga R, Rizo J, Zeng E, Kientz J, Ries R, Otis C, et al. User-centered design of Learn to Quit, a smoking cessation smartphone app for people with serious mental illness. JMIR Serious Games 2018 Jan 16;6(1):e2 [FREE Full text] [doi: 10.2196/games.8881] [Medline: 29339346]

73. Davidson R, Randhawa G, Cash S. Identification of complex health interventions suitable for evaluation: Development and validation of the 8-step scoping framework. JMIR Res Protoc 2019 Mar 05;8(3):e10075 [FREE Full text] [doi: 10.2196/10075] [Medline: 30835240]
74. Wang Y, Fadhil A, Lange J, Reiterer H. Integrating taxonomies into theory-based digital health interventions for behavior change: A holistic framework. JMIR Res Protoc 2019 Jan 15;8(1):e8055 [FREE Full text] [doi: 10.2196/resprot.8055] [Medline: 30664477]
75. Bradway M, Carrion C, Vallespin B, Saadatfard O, Puigdomènech E, Espallargues M, et al. mHealth assessment: Conceptualization of a global framework. JMIR Mhealth Uhealth 2017 May 02;5(5):e60 [FREE Full text] [doi: 10.2196/mhealth.7291] [Medline: 28465282]
76. Shanghai Ranking Consultancy. 2017. Academic ranking of world universities 2017 URL: http://www.shanghairanking.com/ARWU2017.html [accessed 2019-09-06] [WebCite Cache ID 6sjhB63bq]
77. Le Pen C, Priol G, Lilliu H. What criteria for pharmaceuticals reimbursement? An empirical analysis of the evaluation of "medical service rendered" by reimbursable drugs in France. Eur J Health Econ 2003;4(1):30-36. [doi: 10.1007/s10198-002-0145-2] [Medline: 15609166]
78. Hussain M, Al-Haiqi A, Zaidan AA, Zaidan BB, Kiah ML, Anuar NB, et al. The landscape of research on smartphone medical apps: Coherent taxonomy, motivations, open challenges and recommendations. Comput Methods Programs Biomed 2015 Dec;122(3):393-408. [doi: 10.1016/j.cmpb.2015.08.015] [Medline: 26412009]
79. El Amrani L, Oude Engberink A, Ninot G, Hayot M, Carbonnel F. Connected health devices for health care in French general medicine practice: Cross-sectional study. JMIR Mhealth Uhealth 2017 Dec 21;5(12):e193 [FREE Full text] [doi: 10.2196/mhealth.7427] [Medline: 29269336]
80. Yen P, Bakken S. Review of health information technology usability study methodologies. J Am Med Inform Assoc 2012;19(3):413-422 [FREE Full text] [doi: 10.1136/amiajnl-2010-000020] [Medline: 21828224]
81. Michie S, West R, Campbell R, Brown J, Gainforth H. ABC of Behaviour Change Theories: An Essential Resource for Researchers, Policy Makers, and Practitioners. London, UK: Silverback Publishing; 2014.
82. Davis R, Campbell R, Hildon Z, Hobbs L, Michie S. Theories of behaviour and behaviour change across the social and behavioural sciences: A scoping review. Health Psychol Rev 2015;9(3):323-344 [FREE Full text] [doi: 10.1080/17437199.2014.941722] [Medline: 25104107]
83. Czajkowski SM, Lynch MR, Hall KL, Stipelman BA, Haverkos L, Perl H, et al. Transdisciplinary translational behavioral (TDTB) research: Opportunities, barriers, and innovations. Transl Behav Med 2016 Mar;6(1):32-43 [FREE Full text] [doi: 10.1007/s13142-015-0367-3] [Medline: 27012251]
84. Melchert TP. Leaving behind our preparadigmatic past: Professional psychology as a unified clinical science. Am Psychol 2016 Sep;71(6):486-496. [doi: 10.1037/a0040227] [Medline: 27571528]
85. Brunner J, Chuang E, Goldzweig C, Cain CL, Sugar C, Yano EM. User-centered design to improve clinical decision support in primary care. Int J Med Inform 2017 Aug;104:56-64 [FREE Full text] [doi: 10.1016/j.ijmedinf.2017.05.004] [Medline: 28599817]
86. Matthew-Maich N, Harris L, Ploeg J, Markle-Reid M, Valaitis R, Ibrahim S, et al. Designing, implementing, and evaluating mobile health technologies for managing chronic conditions in older adults: A scoping review. JMIR Mhealth Uhealth 2016 Jun 09;4(2):e29 [FREE Full text] [doi: 10.2196/mhealth.5127] [Medline: 27282195]
87. Epstein RM, Street RL. The values and value of patient-centered care. Ann Fam Med 2011;9(2):100-103 [FREE Full text] [doi: 10.1370/afm.1239] [Medline: 21403134]
88. Jaspers MW. A comparison of usability methods for testing interactive health technologies: Methodological aspects and empirical evidence. Int J Med Inform 2009 May;78(5):340-353. [doi: 10.1016/j.ijmedinf.2008.10.002] [Medline: 19046928]
89. Hood L, Heath JR, Phelps ME, Lin B. Systems biology and new technologies enable predictive and preventative medicine. Science 2004 Oct 22;306(5696):640-643. [doi: 10.1126/science.1104635] [Medline: 15499008]
90. Bousquet J, Anto JM, Sterk PJ, Adcock IM, Chung KF, Roca J, et al. Systems medicine and integrated care to combat chronic noncommunicable diseases. Genome Med 2011 Jul 06;3(7):43 [FREE Full text] [doi: 10.1186/gm259] [Medline: 21745417]
91. Balas EA, Boren SA. Managing clinical knowledge for health care improvement. Yearb Med Inform 2000(1):65-70. [Medline: 27699347]
92. Westfall JM, Mold J, Fagnan L. Practice-based research: "Blue Highways" on the NIH roadmap. JAMA 2007 Jan 24;297(4):403-406. [doi: 10.1001/jama.297.4.403] [Medline: 17244837]
93. Riley WT, Glasgow RE, Etheredge L, Abernethy AP. Rapid, responsive, relevant (R3) research: A call for a rapid learning health research enterprise. Clin Transl Med 2013 May 10;2(1):10 [FREE Full text] [doi: 10.1186/2001-1326-2-10] [Medline: 23663660]
94. Bagadiya J. Buddha's Rest Stop. 2017 Jan 03. 151 amazing social media statistics you should know in 2017 URL: https://buddhasreststop.blogspot.com/2017/10/151-amazing-social-media-statistics-you.html [accessed 2017-05-12] [WebCite Cache ID 6qPIbslqL]

95. Neiger BL, Thackeray R, Burton SH, Giraud-Carrier CG, Fagen MC. Evaluating social media's capacity to develop engaged audiences in health promotion settings: Use of Twitter metrics as a case study. Health Promot Pract 2013 Mar;14(2):157-162. [doi: 10.1177/1524839912469378] [Medline: 23271716]
96. Mazzocchi F. Could Big Data be the end of theory in science? A few remarks on the epistemology of data-driven science. EMBO Rep 2015 Oct;16(10):1250-1255 [FREE Full text] [doi: 10.15252/embr.201541001] [Medline: 26358953]
97. Yildirim O, Gottwald M, Schüler P, Michel MC. Opportunities and challenges for drug development: Public-private partnerships, adaptive designs and big data. Front Pharmacol 2016;7:461 [FREE Full text] [doi: 10.3389/fphar.2016.00461] [Medline: 27999543]
98. Tabak RG, Khoong EC, Chambers DA, Brownson RC. Bridging research and practice: Models for dissemination and implementation research. Am J Prev Med 2012 Sep;43(3):337-350 [FREE Full text] [doi: 10.1016/j.amepre.2012.05.024] [Medline: 22898128]
99. Matthews J, Win KT, Oinas-Kukkonen H, Freeman M. Persuasive technology in mobile applications promoting physical activity: A systematic review. J Med Syst 2016 Mar;40(3):72. [doi: 10.1007/s10916-015-0425-x] [Medline: 26748792]
100. Cambon L. Health smart devices and applications...towards a new model of prevention? Eur J Public Health 2017 Jun 01;27(3):390-391 [FREE Full text] [doi: 10.1093/eurpub/ckx019]
101. Moore GF, Audrey S, Barker M, Bond L, Bonell C, Hardeman W, et al. Process evaluation of complex interventions: Medical Research Council guidance. BMJ 2015 Mar 19;350:h1258 [FREE Full text] [doi: 10.1136/bmj.h1258] [Medline: 25791983]

Abbreviations

ACTS: Accelerated Creation-To-Sustainment
BIT: behavioral intervention technology
CeHRes: Center for eHealth Research
CEPS: Plateforme universitaire Collaborative d'Evaluation des programmes de Prévention et de Soins de support
CONSORT: Consolidated Standards of Reporting Trials
CONSORT-EHEALTH: Consolidated Standards of Reporting Trials of Electronic and Mobile HEalth Applications and onLine TeleHealth
DoTTI: Design and develOpment, Testing early iterations, Testing for effectiveness, Integration and implementation
DREAM-GLOBAL: Diagnosing hypeRtension—Engaging Action and Management in Getting LOwer Bp in Aboriginal and LMIC (lower- and middle-income countries)
EVOLVE: Explore Values, Operationalize and Learn, and eValuate Efficacy
FDA: US Food and Drug Administration
IDEAS: Integrate, Design, Assess, and Share
IT: information technology
JITAI: just-in-time adaptive intervention
JMIR: Journal of Medical Internet Research
MOST: Multiphase Optimization STrategy
MRC: Medical Research Council
NHS: National Health Service
NIATx: Network for the Improvement of Addiction Treatment
NIDA: National Institute on Drug Abuse
NIH: National Institutes of Health
ORBIT: Obesity-Related Behavioral Intervention Trials
PRECEDE-PROCEED: Predisposing, Reinforcing, and Enabling Constructs in Educational Diagnosis and Evaluation—Policy, Regulatory, and Organizational Constructs in Educational and Environmental Development
SMS: short message service
TECH: TElehealth in CHronic disease
TUDER: Targeting, Understanding, Designing, Evaluating, and Refining
WHO: World Health Organization

Edited by G Eysenbach; submitted 03.02.19; peer-reviewed by S Bacon, P Rashidi, E Montague; comments to author 23.03.19; revised version received 08.06.19; accepted 02.08.19; published 16.10.19.

Please cite as:
Carbonnel F, Ninot G
Identifying Frameworks for Validation and Monitoring of Consensual Behavioral Intervention Technologies: Narrative Review
J Med Internet Res 2019;21(10):e13606
URL: https://www.jmir.org/2019/10/e13606
doi:10.2196/13606
PMID:31621638

Original Paper

Feasibility, Acceptability, and Preliminary Impacts of Web-Based Patient Education on Patients With Schizophrenia Spectrum Disorder: Quasi-Experimental Cluster Study

Anna Laine[1], RN, MNSc; Maritta Välimäki[1,2], PhD; Virve Pekurinen[1], RN, PhD; Eliisa Löyttyniemi[3], MSc; Mauri Marttunen[4,5], MD, PhD; Minna Anttila[1], PhD

[1]Department of Nursing Science, University of Turku, Turku, Finland
[2]School of Nursing, Hong Kong Polytechnic University, Hong Kong, Hong Kong
[3]Department of Biostatistics, University of Turku, Turku, Finland
[4]Institute of Criminology and Legal Policy, University of Helsinki, Helsinki, Finland
[5]Helsinki University Hospital, Helsinki, Finland

Corresponding Author:
Maritta Välimäki, PhD
Department of Nursing Science
University of Turku
Joukahaisenkatu 3-5
20014 University of Turku
Turku
Finland
Phone: 358 406890546
Email: mava@utu.fi

Abstract

Background: Web-based interventions are promising tools for increasing the understanding of illness and treatment among patients with serious mental disorders.

Objective: This study aimed to test the feasibility and acceptability of a Web-based patient education intervention using a quasi-experimental cluster design to report feedback on patient education sessions and the website used and to report preliminary evidence of the intervention's impact on patients with schizophrenia spectrum disorder.

Methods: A single-blind, parallel, quasi-experimental cluster study over a 6-month period comparing Web-based education (n=33) with a nonequivalent control group (treatment as usual, n=24) for people with schizophrenia spectrum disorder was conducted. Participants (N=57) were recruited from one psychiatric hospital (6 wards). Feasibility was assessed by participants' commitment (refusal rate, dropout rate) to the study. Acceptability was assessed as participants' commitment to the intervention. Patient education sessions and website feedback were assessed by the patients and health care professionals. The preliminary impact of the sessions on patients' self-efficacy, self-esteem, illness cognition, and knowledge level was measured at baseline and follow-ups (8 weeks, 6 months) with self-rated questionnaires.

Results: The refusal rate among patients was high with no statistically significant difference (69% [74/107] in the intervention group, 76% [76/100] in the control group; *P*=.21). The same result was found for the dropout rates (48% [16/33] vs 58% [14/24]; *P*=.46). The acceptability of the intervention was good; 31 participants out of 33 (94%) completed all five sessions. Feedback on the intervention was mainly positive; three out of four subscales of session were rated above the midpoint of 4.0. Feedback on the website was also positive, with a grade of good for content (69%, 20/29 patients; 75%, 21/28 professionals), layout (62%, 18/29 patients; 61%, 17/28 professionals), and usability (62%, 18/29 patients; and 68%, 19/28 professionals). The patients using the intervention had significantly higher scores 6 months after the sessions in self-efficacy (baseline mean 26.12, SD 5.64 vs 6-month mean 29.24, SD 6.05; *P*=.003) and regarding knowledge level about schizophrenia (mean 11.39, SD 4.65 vs 6-month mean 15.06, SD 5.26; *P*=.002), and lower scores in the subscale of helplessness in illness cognition (mean 2.26, SD 0.96 vs 6-month mean 1.85, SD 0.59; *P*=.03). Differences from the control group were not significant. No differences were found in patients' self-esteem or other subscales in illness cognition.

Conclusions: The patients were reluctant to participate in the study and tended to drop out before the follow-ups. Once they had participated, their acceptance of the intervention was high. A more effective recruitment strategy and monitoring method will be needed in future studies. To assess the impact of the intervention, a more rigorous study design with an adequately powered sample size will be used in cooperation with outpatient mental health services.

(J Med Internet Res 2019;21(10):e13073) doi:10.2196/13073

KEYWORDS

mental health; patient education; schizophrenia; feasibility study; internet; information system

Introduction

Background

Schizophrenia, with an age-standardized point prevalence of 0.28% [1], is one of the most severe mental disorders that causes comprehensive impairments to a person's cognitive functioning, requiring long-term treatment [2]. For patients with schizophrenia, nonadherence in treatment is a remarkable problem [3-5] owing to feelings of being stigmatized or self-stigmatized [6]. Continuous symptoms and lack of insight [2] often cause patients to relapse and increase their hospitalizations [7]. Therefore, a variety of psychosocial interventions have been developed for patients with schizophrenia to increase treatment adherence and improve quality of life [2]. Dropout rates in psychosocial interventions have been found to be low [8].

Psychosocial interventions, such as patient education, seem to reduce relapses, readmission, medical nonadherence, and the length of hospital stays of patients with schizophrenia [9]. Various sets of guidelines have also recommended patient education for persons with schizophrenia [2,10]. Patient education includes guidance and information about the illness and how to cope with it. The education can be realized individually or in small groups of patients and their relatives with different kinds of supportive material, such as written procedures, videos, and Web-based programs [9,11]. Previous studies have shown that patients with mental health problems search out health care knowledge on the internet [12-16], which justifies using Web-based interventions in mental health care [12,17].

A variety of Web-based patient education interventions for patients with mental health problems exist, for example, for patients with depression [18-21], bipolar disorder [22], and schizophrenia [23-25]. Studies have shown that Web-based patient education has improved patients' compliance with medication [11] and reduced symptoms [23,26]. Web-based patient education has been found to improve patients' knowledge levels about their illness [23,27]. However, on the basis of Cochrane review by Välimäki et al [11], there is no difference in improvement when compared with other psychoeducational methods used for patients with schizophrenia spectrum disorder.

Objectives

On the basis of the Cochrane review [11], Web-based patient education for patients with schizophrenia is as effective as patient education carried out traditionally, for example, with leaflets, when comparing patients' satisfaction in treatment, even though the number of studies asserting this claim is small and many of them are underpowered. Previous studies have also found that information related to serious mental disorders described on the internet is often low in quality [28]. This finding is concerning because many people with serious mental disorders use the internet as a source of information [16,17]. Therefore, it is critical that online mental health literacy [16] and Web-based interventions [19] are developed for these specific groups, even though it seems that people with mental disorders use the internet less frequently [16,17] compared with the general population [29]. Previous studies have revealed a number of obstacles, such as high refusal and dropout rates [30], although other results have found high acceptability rates [31] and engagement regarding Web- and mobile-based interventions [23,32-35]. On the other hand, Killikelly et al [36] found, in their systematic review, that participants' adherence to Web- and mobile-based interventions ranged between 28% and 100%. Indeed, persons with schizophrenia can be a challenging target group for study recruitment in general [30,37-40]. The aim of this study was to test the feasibility and acceptability of a Web-based educational intervention among persons with schizophrenia using a quasi-experimental study design and to report preliminary evidence of its impact on self-efficacy, self-esteem, knowledge level about schizophrenia, and illness cognition.

Methods

Design and Sample

A single-blind, parallel, quasi-experimental cluster study design was used with a nonequivalent control group for an 8-week timeframe and a 6-month follow-up to evaluate possible short-term and long-term impacts [41] of the intervention. Long-term impacts were included in the assessment as is recommended by the World Health Organization when testing novel digital health interventions [42]. A cluster study design is usable when aiming to avoid information and experience flow between individual participants in intervention and control groups [43,44], as information and experience flow between study groups could have effects on the study results [45,46].

Setting, Eligibility Criteria, and Recruitment

The study was run between May 2015 and May 2016 in 1 psychiatric hospital (6 wards) in Southern Finland. A total of 3 wards caring particularly for patients with schizophrenia spectrum disorder who showed interest in participating in the feasibility study were purposefully invited (a total of 41 beds, 2 closed rehabilitation wards, and 1 closed acute ward treating patients with schizophrenia) and assigned to be the intervention wards. A total of 3 other corresponding wards were then

purposefully selected and invited to join the study owing to their match with the intervention wards (a total of 44 beds, 2 closed rehabilitation wards, and 1 closed acute ward treating patients with schizophrenia) when randomization of participants was not reasonable [46], given the possible information flow between the participants [44].

Patients were eligible to participate if they had been admitted to the study ward during the data collection period, were 18 years old or older, had a primary diagnosis of schizophrenia spectrum disorder (F20–F29, International Classification of Diseases, 10th Revision [ICD-10]) [47], were able to write, read, and speak Finnish, and had volunteered to participate in the study with a written informed consent. Exclusion criteria were an unclear diagnosis, a short hospital period (less than 1 week, which does not allow a proper informed consent process or time to run the intervention), an impaired mental state (assessed by the staff members based on their daily experiences), or a lack of willingness to participate in the study. In addition, if the patient was discharged and rehospitalized during the study period, he or she could only participate in the study once.

One staff member in each study ward with access to hospital medical records acted as a contact person. The contact person was responsible for the patient recruitment and assessment of whether the patients fulfilled the inclusion criteria. The contact person monitored that the rehospitalized patients were not recruited again. The contact person informed the eligible patients about the study (orally and in written format) and further informed the researcher (AL) about the potential participants. Potential participants were informed that they had the possibility to meet the researcher if they wanted more detailed information about the study (voluntary participation, confidentiality, withdrawal without any penalty, or consequences to care) in addition to the contact person's information and other written information. If the patients were willing to participate in the study, they signed 2 informed consent forms [48,49] and provided the baseline data. The completed forms were then sealed in an envelope.

The researcher visited the study wards weekly to ensure that the protocol for patient recruitment was followed and eligible patients were invited to participate in the study. For the follow-up, the instruments were distributed to the patients during their hospital stay (long-term patients) or were sent by post to discharged patients (only with patient permission). The sample size was determined by how many participants were willing and eligible to participate during the 6-month period.

The study was assessed by the Ethics Committee of the Hospital District of Southwest Finland (ETMK:40/1801/2015). The research permission committee of the study organization granted permission for data collection on the study wards. The outcome assessment was carried out according to ethical guidelines [48-50] and the Finnish legislation concerning research [51] and personal data registration [52].

Study Groups

Intervention Group

Patients in the intervention group received Web-based educational intervention by using a health-related website, MentalNet. This website has been originally designed for adult patients with psychosis (ICD-10 codes F20-F29 [47]) to increase their understanding of their illness and its treatment [53]. In its current version, the website targets patients, their relatives, and health care professionals. The website is secured and accessible only with passwords. The website includes the following components: educational material, a discussion forum, and a question and answer column [53,54]. The main component is the educational reading material [23], which includes information for patients with schizophrenia spectrum disorder divided into 5 themes, and tens of accurate and high-quality website links related to each of the 5 topics. There are also tasks for the patients and audio-recorded success stories to increase patients' knowledge about their disorder [9]. The tasks are related to information themes, which were formed according to patients' interests during the patient education sessions. The content of the educational material is described in Table 1. The website has a Health On the Net Code of Conduct certificate as a trustworthy and reliable medical website [55].

The intervention was carried out by health care professionals (nurses, psychologists, and occupational therapists) working in the intervention wards and who were trained to run the patient education sessions and to use the website with the patients [56]. A detailed description of the professionals and Web-based training for professionals to run the patient education sessions are reported elsewhere [56]. The professionals had permanent working positions only in their own intervention or control ward. The intervention included 5 sessions with specific topics based on the information themes of the educational materials on the website: (1) mental disorder, (2) treatment, (3) well-being, (4) patients' rights, and (5) daily life. The professionals scheduled sessions with the patients once per week (each session about 45–60 min); with the exception that if the inpatient stay was planned to be shorter than a week, then a tighter schedule was made (eg, once a day). The professionals also prepared the material needed for each session (computer, internet connection, printer, and a peaceful place). To ensure patient orientation, the order of the themes was not set in advance. Instead, in each session, 1 topic was selected based on the patient's preference. The patient was encouraged to identify any questions or concerns they may have related to the selected topic to discuss with the professional and to use the website to find the answers to his or her questions. The role of the professional was to help the patient focus on questions important to him or her, to help the patient use the website, and to search for information and answers to his or her questions.

Table 1. Content of educational material of MentalNet.

Theme	Information topics	Tasks
Mental disorder	• Impact of the disorder on person with the disorder • Impact of the disorder on caregivers • Different types of psychosis (F20–F29, International Classification of Diseases, 10th Revision) • Depression	• Measures related to mental health
Treatment	• Care and rehabilitation • Life point when treatment is needed • Places where the treatment is realized • Health care professional participating in the treatment • Practical aspects related to treatment • Examinations • Different types of treatment • Restrictive practice	• Time management diary
Well-being	• Mental health • Nutrition • Physical training • Sleep • Hygiene • Intoxicants and smoking • Family and relationships • Sexuality • Work • Education • Spare time • Spiritual well-being	• Healthy eating plate • Physical activity pie • Sleeping diary • Tests about using intoxicants • Social network circle • Relaxing exercise
Patient's rights	• Adequate treatment • Fair treatment • Self-determination • The right to information • Review of documents • Data protection • Involuntary treatment • Patient ombudsman • Patient injury • Laws related to patient's rights	• Questions about patient's rights
Daily life	• Economic support • A guardian • Living • Support for taking care of home • Return to work • Support from fellow man	• Test related to instrumental activities of daily living

Control Group

Patients in the control group continued their treatment with care as usual. They did not receive any psychoeducational intervention provided by the researcher or have access to the website.

Outcomes and Assessment Instruments

Primary Outcomes

Feasibility

Feasibility was assessed by patient refusal of the study (yes, no), participation in the follow-ups (yes, no), and whether they dropped out of the study (attrition rate).

Acceptability

Acceptability was assessed by the patients in the intervention group. The number of patients participating in all possible sessions (5 sessions per participant) and the total amount of sessions were calculated.

Feedback of the Patient Education Sessions

The one-on-one patient education sessions were assessed by both the patient and the professional using the Finnish translation of the Session Evaluation Questionnaire (Form 5) (SEQ) [57]. SEQ is a self-rating instrument originally designed to measure psychotherapy and counselling sessions and to measure participants' (client and/or therapist) moods after the session. The SEQ includes 21 bipolar adjective items divided into 2 parts: 11 items about the session itself (1 global item *bad–good*,

5 subscale items for *depth*, and 5 subscale items for *smoothness*) and 10 items about participants' moods after the session (5 subscale items for *positivity* and 5 subscale items for *arousal*), such as "This session was *valuable-worthless*, *easy-difficult*" or "Right now I feel *happy-sad*, *angry-pleased*." The range of the scale is 1 to 7, with a midpoint of 4.00. The mean scores of each subscale items are calculated to form a subscale score, and mean values above the midpoint of 4.00 are considered to be a positive evaluation of the session [58]. In previous studies, the internal consistency of the instrument has been found to be relatively good (Cronbach alpha .63-.93) [58]. In our study, Cronbach alphas ranged between .34 and .84 (patients: depth .67, smoothness .73, positivity .73, arousal .39; and professionals: depth .81, smoothness .75 positivity .84, arousal .34). The instrument was translated into Finnish with a back-translation method [59,60] using an independent professional translator and the original developer of the questionnaire.

Feedback of the Website

Feedback about the MentalNet website was collected from the patients and the professionals, after all 5 sessions were completed, using a 5-point Likert scale (very good–very poor) and with the possibility to give written feedback. The feedback targeted the content, layout, and usability of the website to ensure that the website is usable and meets the needs of its users [61,62].

Secondary Outcomes at Baseline, 8 Weeks, and 6-Month Follow-Up

Self-Efficacy

The General Self-Efficacy Scale (GSE) [63] is a widely used self-rating instrument designed to measure the general sense of perceived self-efficacy in different types of difficult life events. The instrument contains 10 items, and its responses are in the form of a 4-point Likert scale. The sum score of the responses ranges from 10 to 40; a higher score represents greater sense of self-efficacy. In a study by Scholz et al [64], the psychometric properties of the GSE were examined in 25 countries, and the Cronbach alpha varied from .75 to .91. In our study, the Cronbach alpha value varied between .92 and .96.

Self-Esteem

The Rosenberg Self-Esteem Scale (SES) [65] is a self-rating instrument designed to measure overall self-esteem. It includes 10 items with a 4-point Likert scale. The sum score of the answers ranged from 10 to 40; higher scores indicate higher self-esteem. The SES has been translated into at least 28 languages and is widely used in many countries. In a review by Schmitt and Allik [66], the data of SES from 53 countries were compared. Internal consistency was found to be good (Cronbach alpha .80, range .45 to .90). In our study, the Cronbach alpha value varied between .83 and .90.

Illness Cognition

The Illness Cognition Questionnaire [67] is a self-rating instrument designed to measure illness cognition of people with chronic illnesses as how they perceive and think about their illness. The participants are asked to assess to what extent they assess with 18 statements of the instrument by using a 4-point Likert scale (1=not at all, 2=somewhat, 3=to a large extent, and 4=completely). The instrument consists of 3 subscales (6 items each) measuring a basic set of illness cognitions that includes both unfavorable (negative) and favorable (positive) ways of adjusting to chronic disease: *helplessness* (with negative perspective, eg, "My illness limits me in everything that is important to me."), *acceptance* (with positive perspective, eg, "I can handle the problems related to my illness."), and *perceived benefits* (with a positive perspective, eg, "I have learned a great deal from my illness."). Internal consistency has proven to be adequate when using the instrument, for example, with patients with chronic pain (Cronbach alpha for helplessness .88, acceptance .91, and perceived benefits .83) and chronic fatigue (Cronbach alpha for helplessness .83, acceptance .90, and perceived benefits .81) [68]. In this study, the Cronbach alpha varied between .77 and .97 for helplessness, .74 and .91 for acceptance, and .61 and .91 for perceived benefits. The instrument was translated into Finnish with a back-translation method [59,60] using an independent professional translator and the original developer of the questionnaire.

Knowledge Level

Knowledge about Schizophrenia Questionnaire [69] is a self-rating instrument designed to measure the knowledge of patients with schizophrenia about their illness and its management. The instrument is a multiple-choice test with 25 items with themes as follows: diagnosis, frequency, etiology, progress and prognosis of illness, medication and its side effects, drug-free treatments, stress, and legislation. The respondent is given 1 point for each correct answer, and the sum score ranges between 0 and 25; a higher score represents a high knowledge level. Internal consistency of the instrument is proven to be good (Cronbach alpha .75 [69]). In this study, the Cronbach alpha varied between .71 and .81. The instrument was translated into Finnish with a back-translation method [59,60] using an independent professional translator and the original developer of the questionnaire. A cultural modification was made for question number 20 to fit the Finnish health care system and the Mental Health Act about involuntary treatment [70].

Sociodemographic Information and Internet Use at Baseline

Information about the patients' age, gender, age at first contact with psychiatric services, attitudes toward computers or the internet, and their computer or internet skills was collected. Attitudes toward computers or the internet and their computer or internet skills were assessed with a 5-point Likert scale (*Your attitude toward using the computer/internet is [1=very positive to 5=very negative*] and *Your computer/internet skills are [1=very good to 5=poor*]). Moreover, the adapted instrument [16] by Choi and DiNitto [71] was used to describe participants' internet use and purpose of internet use. The options *communicate with health professionals about health-related issues* and *communicate with other users about health-related issues* were added [16] to the original questions about internet use [71].

Data Analysis

Descriptive statistics were used for numerical variables with a median, mean, and standard deviation (SD), whereas categorical variables are reported with counts and percentages, and sum scores for each scale were calculated. Feedback from the patients and the professionals regarding the MentalNet website were compared using a Chi-square test (x^2). Data regarding self-efficacy, self-esteem, illness cognition, and knowledge level were analyzed with hierarchical linear mixed models, allowing subjects to have missing values. The analysis included all 3 time points (baseline, 8 weeks, and 6 months). The model was adjusted by age, gender, and group of participation. One main interest was to focus on whether the mean change between time points differed among the groups. A compound symmetry covariance structure was used for repeated measures. A Cronbach alpha was calculated for all questionnaires. These statistical tests were performed as 2-tailed, with a significance level set at .05. The analyses were performed using the SAS System, version 9.4 for Windows (SAS Institute Inc.).

Cohen *d* was calculated between baseline and 6 months to find out the effect size of the intervention on patients' self-efficacy, self-esteem, illness cognition, and knowledge level. Based on Cohen, 0.2 is considered a small, 0.5 is a medium, and 0.8 is a large effect size [72]. The analysis was performed using an online calculator [73].

Results

Sociodemographic Information

In both the intervention and control groups, the mean age of the patients was approximately 41 years (range 20–66 in the intervention group and 23–69 in the control group). Their age when they first accessed mental health care services was around 24 years (range 8–47 in the intervention group and 3–68 in the control group). The proportion of women in the intervention group was twice as much as it was in the control group.

Although the patients' internet skills varied, about 70% reported that they used the internet and about 40% had good internet skills (very good or good). A clear majority had positive attitudes toward computers. Patients used the internet most often for searching for knowledge other than health information, emailing, banking, reading news or books, and social media. Background information about patients and their use of the internet are presented in more detail in Table 2.

Table 2. Background information of the patients.

Patient's information	Intervention group (N=33)	Control group (N=24)
Age (years), mean (SD)	42 (14.13)	41 (12.90)
Age (years) when first receiving mental health care, mean (SD)	24 (9.30)	24 (12.60)
Gender, n (%)		
Female	18 (55)	6 (25)
Male	15 (45)	18 (75)
Use of internet, n (%)		
Never user	6 (18)	4 (17)
Previous user	4 (12)	3 (13)
Current user	23 (70)	17 (71)
Computer/internet skills, n (%)		
Very good	3 (9)	4 (17)
Good	10 (30)	6 (25)
Neither good nor bad	6 (18)	7 (29)
Fairly poor	1 (3)	2 (8)
Poor	13 (39)	5 (21)
Attitudes toward computers/internet, n (%)		
Very positive	8 (24)	8 (33)
Positive	15 (45)	10 (42)
Neither positive nor negative	7 (21)	4 (17)
Negative	3 (9)	2 (8)
Very negative	0 (0)	0 (0)
Purpose of the internet use, n (%)		
Research information about other topics of interest	15 (45)	16 (67)
Send/receive email	14 (42)	15 (63)
Watch videos	11 (33)	15 (63)
Do banking online and/or pay bills	11 (33)	14 (58)
Read newspapers, magazines, and books online	7 (21)	16 (67)
Use social networking or dating site	9 (27)	14 (58)
Research health-related information	10 (30)	10 (42)
Play games online	8 (24)	11 (46)
Buy products online	5 (15)	11 (46)
Communication with others	5 (15)	3 (13)
Other	3 (9)	6 (24)
Communication with health professionals	2 (6)	3 (13)

Feasibility

A flow diagram of participating patients is presented in Figure 1. During the data collection, 303 patients were assessed for eligibility. A total of 213 patients were invited to participate in the study and 150 of them refused. The refusal rate between the intervention and control groups was not statistically significant (69% [74/107] in the intervention group vs 76% [76/100] in the control group, *P*=.21). At baseline, out of the allocated patients, 58% (33/57) were in the intervention group and 42% (24/57) in the control group. After baseline, 33% (11/33) dropped out of the intervention group and 46% (11/24) dropped out of the control group. In total, 8 patients dropped out after the first follow-up: 5 from the intervention group and 3 from the control group. This left us with a total of 63% (17/27) patients in the intervention group and 37% (10/27) in the control group. The difference in the number of dropouts between the intervention and control groups was not statistically significant (48% [16/33] vs 58% [14/24]; *P*=.46). No statistically significant differences between completers and dropouts were found regarding their

gender, age, or age at first time of received mental health care (see Table 3).

Acceptability of the Intervention

It was planned that each study participant (N=33) would have 5 intervention sessions. Out of 33 participants, 31 patients (94%) had all 5 sessions of the intervention. Out of the remaining 2 patients, 1 patient completed 3 sessions and 1 patient 1 session. Altogether, 159 (96%) sessions out of 165 planned sessions were realized.

Feedback of the Patient Education Sessions

Each of the 5 patient education sessions were evaluated by the patients and the professionals directly afterward. The means of the global item *bad-good* and the subscales *depth*, *smoothness*, and *positivity* were evaluated to be above the midpoint of 4.0 (with a range of 1–7) by both the patients and professionals. Patients' evaluations were more positive with a statistically significant difference in the global item *bad-good* (*P*=.02) and in the subscale *depth* (*P*=.04), and professionals' evaluations were more positive with a statistically significant difference in the subscale *positivity* (*P*=.03) when evaluations were compared with each other (see Table 4).

Figure 1. Flow diagram of patients.

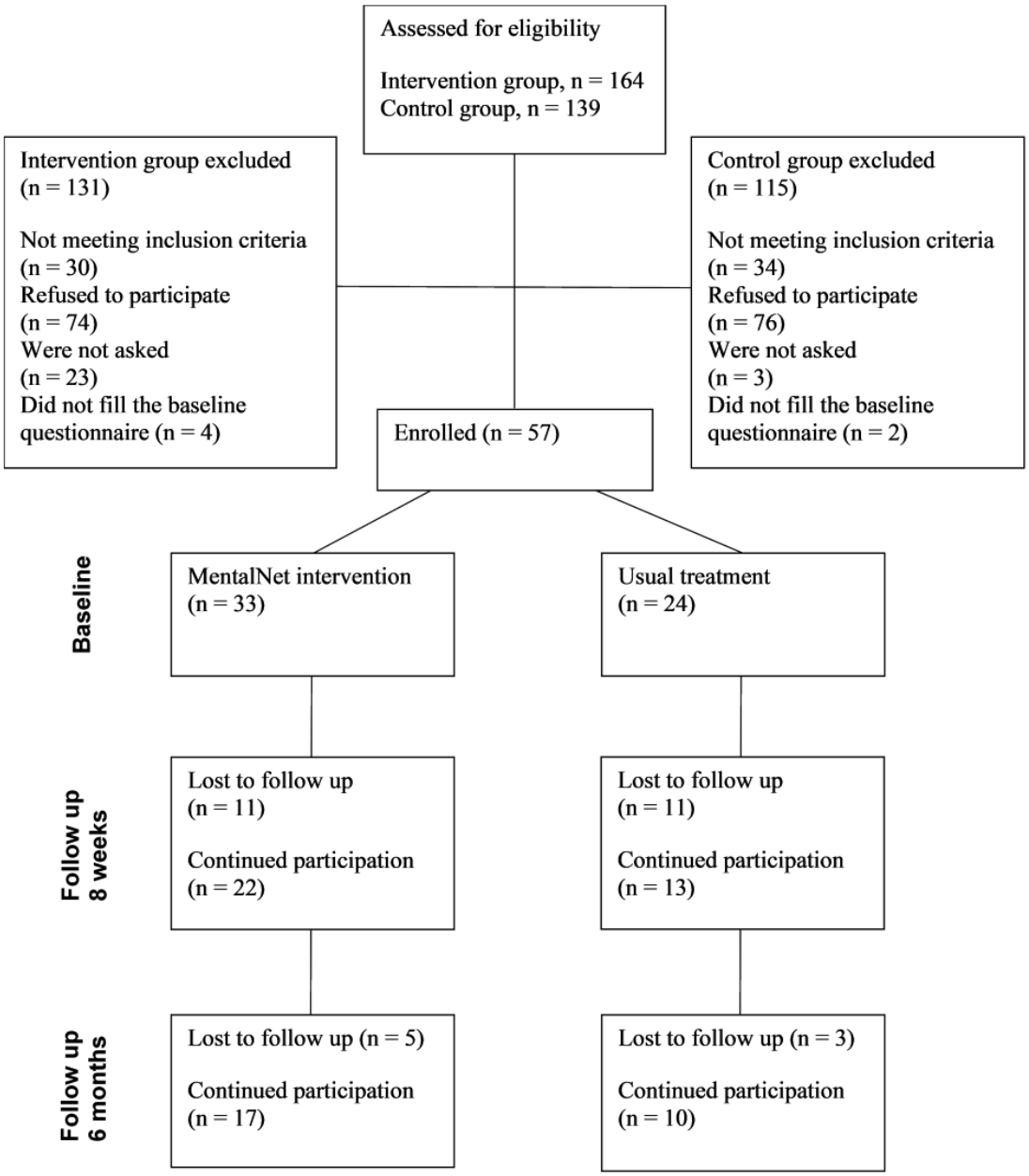


Table 3. Demographic characteristics of patients who dropped out and completed the study.

Demographic characteristics	Dropouts (N=30)	Completers (N=27)	*P* value
Age (years), mean (SD)	40 (14.33)	43 (12.61)	.30
Age (years) when first received mental health care, mean (SD)	24 (10.08)	23 (11.49)	.27
Gender (female), n (%)	11 (37)	14 (52)	.38

Table 4. Session evaluation of Web-based patient education meetings by patients and health care professionals combined from all 5 sessions (n=number of sessions) based on dimensions of Session Evaluation Questionnaire [58].

Dimension	Patients (N=33)			HCPs[a] (N=33)			*P* value
	n	Mean (SD)	Median	n	Mean (SD)	Median	
Bad-good	154	5.51 (1.49)	6.00	152	5.13 (1.40)	6.00	.02[b]
Depth	150	4.57 (1.15)	4.60	146	4.31 (0.96)	4.20	.04[b]
Smoothness	151	4.99 (1.12)	5.00	146	5.00 (0.85)	5.00	.98
Positivity	149	5.27 (1.17)	5.40	138	5.54 (0.82)	5.80	.03[b]
Arousal	153	3.47 (0.94)	3.60	138	3.33 (0.58)	3.40	.12

[a]HCPs: health care professionals.

[b]Statistically significant difference.

Feedback of the Website

The patients and the professionals gave their feedback on the website. Most (≥65%) of the patients and the professionals responded that the content, layout, and usability of MentalNet was good or very good. The numerical feedback is presented in more detail in Table 5.

In the written feedback, patients expressed thanks for the opportunity to participate in the intervention, which they had found meaningful. Patients were satisfied with the comprehensive and good content of the website. They had been able to get important information that had helped them to understand their situation and would be valuable in the future. Critical feedback from patients was related to tasks that could be too difficult if the patient lacked the computer skills and if the links were not working.

The professionals offered, in their insight, that the content of the website included comprehensive information with good themes. In their opinion, using the website gave a structure for patient education. The professionals felt that these themes with new or iterated information for patients was important to go through, and the professionals expressed that they would use the website again in the future. The professionals were also able to gain new information about patients. This information could be used for better care of the patients in the future. In their opinion, the website was useful and easy to use with those patients who were enthusiastic about the intervention, able to use the website, and willing to find information independently. On the other hand, some professionals felt that going through the website was useless and it was hard to get patients interested in using it. The professionals gave critical feedback about the content and layout of the website. In their opinion, some pages included too much information, and it was therefore hard for patients to follow. The professionals noticed that some links were not working, and some felt that there were too many links. The layout of the website was considered old fashioned, and the professionals proposed that it should be updated and clarified. They also recommended that more tasks especially related to patients with psychosis should be added, which could increase patients' illness recognition.

Table 5. Feedback on the website.

Dimension of the feedback	Patients (N=29), n (%)	HCPs[a] (N=28), n (%)
Content of the website		
Very good	5 (17)	3 (11)
Good	20 (69)	21 (75)
Not good or poor	4 (14)	4 (14)
Poor	0 (0)	0 (0)
Very poor	0 (0)	0 (0)
Layout of the website		
Very good	2 (7)	1 (4)
Good	18 (62)	17 (61)
Not good or poor	6 (21)	7 (25)
Poor	3 (10)	3 (11)
Very poor	0 (0)	0 (0)
Usability of the website in patient education		
Very good	4 (14)	2 (7)
Good	18 (62)	19 (68)
Not good or poor	6 (21)	3 (11)
Poor	0 (0)	4 (14)
Very poor	0 (0)	0 (0)

[a]HCPs: Health care professionals.

The Preliminary Impact of the Web-Based Course on Patients' Self-Efficacy, Self-Esteem, Illness Cognition, and Knowledge About Schizophrenia

The preliminary impact of the Web-based intervention was measured at 3 time points (baseline, 8 weeks, and 6 months). Multimedia Appendix 1 shows the results of the hierarchical linear mixed models for repeated measures. Overall, there were no significant differences in time-by-group interaction with any instrument measured in this study. In a more detailed examination, we found that patients' self-efficacy scores increased in the intervention group (at baseline: mean 26.12, SD 5.64), after the intervention (8 weeks: mean 26.50, SD 7.20), and after 6 months (mean 29.24, SD 6.05). The self-efficacy scores also increased in the control group (at baseline: mean 27.26, SD 9.36), after 8 weeks (mean 31.69, SD 6.60) but not after 6 months (mean 30.80, SD 6.41). The change between baseline and the 6-month follow-up was statistically significant in the intervention group (P=.003) but not in the control group. Also, the effect size (d=0.53) refers to medium effect in the intervention group and small in the control group (d=0.44). There were no statistical differences in patients' self-esteem in either group during the 6-month study period.

The subscale *helplessness* in illness cognition decreased in the intervention group (at baseline: mean 2.26, SD 0.96) after the intervention (mean 2.11, SD 0.72) and after 6 months (mean 1.85, SD 0.59). The change between baseline and 6 months was statistically significant (P=.03). Also, the effect size (d=0.51) refers to medium effect in helplessness in the intervention group and small in the control group (d=0.17). The change in the control group was not significant, and there were no significant differences in the intervention group or the control group regarding the other subscales. Further, the knowledge level of patients in the intervention group increased (at baseline: mean 11.39, SD 4.65) after the intervention (mean 12.50, SD 5.26) and after 6 months (mean 15.06, SD 5.26). The change between baseline and 6 months was statistically significant (P=.002). The knowledge level of the control group stayed stable. Also, the effect size (d=0.74) refers to medium effect in the intervention group and small in the control group (d=0.07).

Discussion

Principal Findings

The aim of our study was to test the feasibility and acceptability of a Web-based patient education intervention and to report preliminary evidence of its impact on patients with schizophrenia. The feasibility was assessed by participants' commitment to the study. We found that, in general, during the recruitment period, patients' refusal rates were high (69% in the intervention group and 76% in the control group), which is congruent with previous studies on patients with schizophrenia [30,37-39] even though variety exists [36]. The reasons for refusal were not asked about, as study participation was voluntary, based on ethical guidelines [50], and we did not have consent to collect that information. However, the researcher visited the study wards regularly and discussed practical issues of the study with the professionals and patients. Some patients may have been concerned about the aim of the study and were

therefore suspicious of it. For example, they might have been apprehensive about the confidentiality related to the research [74-76]. Previous studies have found that suspiciousness of studies and/or researchers can be one of the reasons for a high refusal rate, as it is known to be one of the symptoms of schizophrenia when the patient is in psychosis [2,74,77,78]. Even though many patients have had positive experiences with technology use [79], some patients might be afraid to engage themselves with such a study if they have difficulties with concentration [80] or they think their inpatient stay will only last for a short period of time. Moreover, some patients also think that they cannot participate in a study because they do not have schizophrenia (also [77,81]), a diagnosis that can be difficult and time consuming for some patients to accept [2]. According to a study by Woodall et al [23], the timing of asking for consent can also affect a patient's decision to participate or not. In our study, recruitment often took place soon after a patient had been admitted to the hospital. It is therefore possible that this was too soon for some patients, because in the beginning of care, illness may be in an acute phase. There might also be patients who are wary of trying new types of treatment and would rather concentrate on proven traditional methods without any extra distraction.

High refusal rates can also be the result of professionals' involvement, which was found to be the case in a study by Jørgensen et al [82]. Professionals may question whether their patients are too severely ill to participate or to make the decision to participate [81,82]. Therefore, in our study, the researcher reminded the professionals about recruitment and encouraged them to invite patients to participate when visiting the study wards. It is also possible that some patients refused because they did not get enough detailed information about the study. Study recruitment and studies can sometimes be seen as an extra task not related to basic nursing care. Another reason for refusals could be that there were some inconsistencies in the study recruitment; out of 303 eligible patients, there were a total of 26 patients whose willingness to participate was not asked.

The attrition rate was high in both groups (see also Kannisto et al [30]). Participation was voluntary, and reasons for deciding to dropout were not questioned, based on ethical guidelines and principles [48,49]. However, we can suggest some explanations for the dropouts that did occur. For example, some participants were discharged from the hospital before the follow-up, which made contact with them more challenging than it was with those who were able to participate in the follow-up during their hospital stay. Even though the researcher tried to contact all of them by phone and by sending them the follow-up questionnaires at home, not all had a phone, some did not answer the phone, a phone number was not in use, and 1 had a call blocker set up. In addition, some patients did not have a permanent home address or their address was unknown. Therefore, some of the reasons for dropout can be said to be because of the service system, which makes it difficult to maintain contact with the patients. However, other reasons for dropout may relate to a patient's mental status and include lack of interest and tiredness. Patients can express interest in having something to do in the hospital, but then not actually have the time to participate [75,77]. This notion has been mentioned in a study by Furimsky et al [81], when patients with psychosis did not want to continue the study because it took up their time or they felt that their well-being had increased during the follow-up period and therefore did not think that they would benefit from participating any more. It is also possible that, in our study, patients were not willing to think about their disorder anymore or thought that, because of improved well-being and being discharged from the hospital, they were not suitable for the study anymore.

On the other hand, out of 33 participants, 31 (94%) finished all 5 intervention sessions in the patient education. This result strengthens the results of earlier studies where patients with schizophrenia spectrum disorders have engaged in Web-based patient education [23,32]. Further, Villeneuve et al [8] showed that the dropout rates from psychosocial treatments are low (13%) among patients with schizophrenia spectrum disorders. Our finding may indicate that, as soon as the participants were engaged in the study, they accepted and wanted to join the sessions. This result is supported by the evaluation of each session. The patients and the professionals alike gave positive feedback on the patient education sessions; the mean scores of the SEQ were over the midpoint of 4.0 regarding the global item *bad-good* and the subscales *depth*, *smoothness*, and *positivity*. When the session evaluations of the patients and the professionals were compared, the patients were shown to be even more convinced than the professionals that the sessions were good and deep. On the other hand, according to the professionals' evaluations, their moods after the sessions were more positive than the moods of the patients. In a study by Kivlighan et al [83], clients' evaluations were similarly more positive in depth and smoothness and therapists' evaluations were stronger in positivity. This does not seem to be a general trend, however, when the evaluations between therapists and clients have varied [58,84].

In this study, to ensure the patient orientation of the intervention, the participants decided the order of the themes of their sessions. Therefore, it is not possible to directly compare the sessions and evaluate the differences between them. However, it is possible that the patients chose the theme most important to them as the theme for their first sessions and also evaluated those sessions with the highest scores. Patients might have been especially interested in, for example, patients' rights if they were unaware of their own illness or had been involuntarily admitted. The topic *mental disorder* could have been uninteresting or unimportant to a patient if they did not consider themselves to be suffering from the disorder. Therefore, in the future, more attention should be given to the order and evaluations of the sessions. Notably, results of the subscale *arousal* need to be interpreted with caution because of the heterogeneity of the items in the scale. The feedback on the website was also positive regarding its content, layout, and usability. The positive perception of the website may help its use in the future.

Our preliminary results did not find any statistically significant differences between the intervention and control groups, indicating that the intervention did not have an impact on participants' self-efficacy, self-esteem, illness cognition, or knowledge level any more than care as usual. On the other hand,

we found that patients' self-efficacy and knowledge levels between baseline and 6 months in the intervention group improved, while no improvement was seen in the control group. This may indicate positive outcomes of our educational intervention, although, most probably due to insufficient statistical power, not statistically different compared with care as usual. The results might also suggest that discharging a patient from the hospital between baseline and the follow-up may be a confounding factor, which could explain the positive course of self-efficacy and mental health after discharge from psychiatric hospital [85,86]. Further, we found that the knowledge of the participants improved after the Web-based patient education intervention, which has been supported in earlier studies [23].

The subscale *helplessness* in illness cognition decreased significantly in the intervention group, but not in the control group, between baseline and 6 months. In addition, self-esteem increased in both study groups, but the difference was not significant. There is little previous knowledge of how Web-based patient education impacts patients' illness cognition and self-esteem, and therefore, more research with bigger sample sizes is needed.

Most of the patients (70% of the intervention group and 71% of the control group) used the internet, and only a few (8% of the intervention group and 9% of the control group) had negative attitudes about the internet. These findings are similar to earlier studies concerning internet use [17] and attitudes toward it [16] among Finnish patients with schizophrenia. The patients used the internet for diverse purposes, just as the rest of internet users in the Finnish population [29]. However, on the basis of the patients' responses, even if their attitude toward the internet was mostly positive, their internet skills varied. We may therefore ask whether patients have enough confidence to use information technology as part of their treatment. To increase patients' confidence, we should improve their skills with training by using the internet with them to ensure their computer and internet skills. Another concern is whether professionals are really ready to apply information technology in daily treatment practices [87,88]. Still, most patients and professionals found the content, layout, and usability of the website to be good.

In the control group, there were over twice as many male participants as female participants (18 vs 6). In the intervention group, the difference was minor (18 vs 15). In Finland, 48% of patients in inpatient psychiatric care are male, even though the number of male patients is greater than that of female patients among patients of working age [89]. Therefore, we may ask if female patients were more interested in participating in the intervention group than in the control group. It is therefore possible that female patients were more interested in participating in the intervention than male patients. These are important aspects when planning future interventions and studying them on a larger scale.

Limitations

There are some limitations to this study. First, the patient attrition rate was high; we lost many participants after their discharge. Therefore, a hierarchical linear mixed model was used for the data analysis of the preliminary impacts because it allows subjects to have missing values, and we were able to use all data including subjects with missing values. Still, more detailed and systematic information about the number of these patients and reasons for not being able to contact them would be beneficial to include in future studies to get more specific information about attrition when planning new studies with bigger samples and enhanced statistical power. We now also have a hint that this type of Web-based education as a patient intervention should be designed in collaboration with inpatient and outpatient psychiatric services. Second, the sample size was small, which means that the study did not have enough statistical power to demonstrate effects of the intervention or differences between the study groups significantly and trustworthily. Owing to the underpowered sample size, study results must be considered carefully. Possible impacts of the intervention should be measured in a future study with a sufficient sample size that could compare the results of patients in rehabilitation and acute wards. Also, the instruments used for self-efficacy and self-esteem are generic [63,65] and not specific for patients with mental disorders, which might be another explanation for why improvements were not detected. Third, the last follow-up measurement was 6 months from the baseline. It is uncertain whether the changes were only due to the intervention or if they had been influenced by the time elapsed with other possible factors affecting the results [90]. Therefore, studies with a more robust design are needed in the future. In this task, we received many important improvement strategies that can be used to develop this kind of study. Fourth, owing to the quasi-experimental study design, participants were not randomized into the groups. Selection bias was minimized with baseline measurements, where we did not find differences between the groups [46]. Fifth, patients in the control group continued their treatment with care as usual without having any psychoeducational intervention from the research. Psychoeducation is, however, part of the schizophrenia care guideline in Finland [10], and it is possible that patients in the control group received some information from the professionals about their illness. Finally, the Cronbach alpha of the SEQ was low, especially in the subscale *arousal* calculated from both groups (patients and professionals), indicating that not all the participants understood the subscale items the same way. Therefore, the results regarding this subscale should be interpreted with caution. In future studies, the possibility of only using 3 subscales could be discussed [91].

Conclusions

Although the feasibility of the intervention was poor given high refusal and dropout rates, we found that the acceptability of the intervention was good in terms of session completion. Also, the feedback on the sessions and website was positive. Furthermore, the Web-based intervention showed promising impacts on patients' self-efficacy and knowledge levels in the intervention group. There is, therefore, potential to empower patients in the use of Web-based patient education, in terms of increased self-efficacy and illness cognition. However, a future study of this target group will require more effective strategies for recruitment, motivating patients in participation, and monitoring to decrease dropouts, especially when patients leave the hospital during the follow-up period.

Implications

This study revealed important information about numbers of refusal and dropout and showed that more effective monitoring is needed to ensure that all possible participants are screened for eligibility and asked to participate in studies such as this. In the recruitment process, it means, for example, closer cooperation with the contact persons in the study wards during the research, but also in advance, to ensure that the recruitment process is implemented smoothly from the very beginning. Furthermore, the schedule for the researcher's visits to the study wards could be made available in advance to staff members, possible participants, and patients who are already recruited. They would then have easy access to the researcher for asking questions before and during the research, which could improve recruitment and decrease the number of dropouts. In future studies, to increase knowledge about the subject, it might also be good to include a voluntary question asking the reasons for any refusals. Most of the patients had positive attitudes toward the internet and computers and most of them use the internet, but their skills in doing so need improving. This issue provides potential for the professionals and puts them in an important role. Therefore, it is important that professionals have the desire and the resources to support and improve patients' internet skills to include Web-based patient education in everyday patient treatment.

Acknowledgments

The authors thank the study hospital and patients and professionals who participated in this study. The authors also thank the copyright holders of the instruments for granting the permission to use, and translate, when necessary, the instruments in this study. The study was funded by the University of Turku (26002441).

Conflicts of Interest

None declared.

Multimedia Appendix 1
Self-efficacy, self-esteem, illness cognition, and knowledge level at baseline, 8 weeks, and 6 months.
[PDF File (Adobe PDF File), 166 KB - jmir_v21i10e13073_app1.pdf]

References

1. Charlson FJ, Ferrari AJ, Santomauro DF, Diminic S, Stockings E, Scott JG, et al. Global epidemiology and burden of schizophrenia: findings from the global burden of disease study 2016. Schizophr Bull 2018 Oct 17;44(6):1195-1203. [doi: 10.1093/schbul/sby058] [Medline: 29762765]
2. NICE | The National Institute for Health and Care Excellence. 2014. Psychosis and Schizophrenia in Adults: Prevention and Management URL: https://www.nice.org.uk/guidance/cg178/ [accessed 2019-08-30]
3. Lieberman JA, Stroup TS, McEvoy JP, Swartz MS, Rosenheck RA, Perkins DO, Clinical Antipsychotic Trials of Intervention Effectiveness (CATIE) Investigators. Effectiveness of antipsychotic drugs in patients with chronic schizophrenia. N Engl J Med 2005 Sep 22;353(12):1209-1223. [doi: 10.1056/NEJMoa051688] [Medline: 16172203]
4. Higashi K, Medic G, Littlewood KJ, Diez T, Granström O, de Hert M. Medication adherence in schizophrenia: factors influencing adherence and consequences of nonadherence, a systematic literature review. Ther Adv Psychopharmacol 2013 Aug;3(4):200-218 [FREE Full text] [doi: 10.1177/2045125312474019] [Medline: 24167693]
5. Velligan DI, Weiden PJ, Sajatovic M, Scott J, Carpenter D, Ross R, Expert Consensus Panel on Adherence Problems in Serious and Persistent Mental Illness. The expert consensus guideline series: adherence problems in patients with serious and persistent mental illness. J Clin Psychiatry 2009;70(Suppl 4):1-46; quiz 47 [FREE Full text] [Medline: 19686636]
6. Gerlinger G, Hauser M, de Hert M, Lacluyse K, Wampers M, Correll CU. Personal stigma in schizophrenia spectrum disorders: a systematic review of prevalence rates, correlates, impact and interventions. World Psychiatry 2013 Jun;12(2):155-164 [FREE Full text] [doi: 10.1002/wps.20040] [Medline: 23737425]
7. Morken G, Widen JH, Grawe RW. Non-adherence to antipsychotic medication, relapse and rehospitalisation in recent-onset schizophrenia. BMC Psychiatry 2008 Apr 30;8:32 [FREE Full text] [doi: 10.1186/1471-244X-8-32] [Medline: 18447935]
8. Villeneuve K, Potvin S, Lesage A, Nicole L. Meta-analysis of rates of drop-out from psychosocial treatment among persons with schizophrenia spectrum disorder. Schizophr Res 2010 Aug;121(1-3):266-270. [doi: 10.1016/j.schres.2010.04.003] [Medline: 20452749]
9. Xia J, Merinder LB, Belgamwar MR. Psychoeducation for schizophrenia. Cochrane Database Syst Rev 2011 Jun 15(6):CD002831 [FREE Full text] [doi: 10.1002/14651858.CD002831.pub2] [Medline: 21678337]
10. Salokangas RK, Tuominen L, Koponen H, Laukkala T, Oksanen J, Pirkola S, et al. Käypä Hoito. 2013. Schizophrenia: Current Care Guidelines URL: https://www.kaypahoito.fi/en/ccs00038 [accessed 2019-08-30]
11. Välimäki M, Hätönen H, Lahti M, Kuosmanen L, Adams CE. Information and communication technology in patient education and support for people with schizophrenia. Cochrane Database Syst Rev 2012 Oct 17;10:CD007198. [doi: 10.1002/14651858.CD007198.pub2] [Medline: 23076932]

12. Kalckreuth S, Trefflich F, Rummel-Kluge C. Mental health related internet use among psychiatric patients: a cross-sectional analysis. BMC Psychiatry 2014 Dec 24;14:368 [FREE Full text] [doi: 10.1186/s12888-014-0368-7] [Medline: 25599722]
13. Conell J, Bauer R, Glenn T, Alda M, Ardau R, Baune BT, et al. Online information seeking by patients with bipolar disorder: results from an international multisite survey. Int J Bipolar Disord 2016 Dec;4(1):17 [FREE Full text] [doi: 10.1186/s40345-016-0058-0] [Medline: 27552813]
14. Baup H, Verdoux H. Frequency and pattern of internet use in patients with schizophrenia or bipolar disorders seeking medical information. Psychiatry Res 2017 Jan;247:152-154. [doi: 10.1016/j.psychres.2016.11.028] [Medline: 27893996]
15. Villani M, Kovess-Masfety V. How do people experiencing schizophrenia spectrum disorders or other psychotic disorders use the internet to get information on their mental health? Literature review and recommendations. JMIR Ment Health 2017 Jan 3;4(1):e1 [FREE Full text] [doi: 10.2196/mental.5946] [Medline: 28049620]
16. Athanasopoulou C, Välimäki M, Koutra K, Löttyniemi E, Bertsias A, Basta M, et al. Internet use, ehealth literacy and attitudes toward computer/internet among people with schizophrenia spectrum disorders: a cross-sectional study in two distant European regions. BMC Med Inform Decis Mak 2017 Sep 20;17(1):136 [FREE Full text] [doi: 10.1186/s12911-017-0531-4] [Medline: 28931385]
17. Välimäki M, Kuosmanen L, Hätönen H, Koivunen M, Pitkänen A, Athanasopoulou C, et al. Connectivity to computers and the internet among patients with schizophrenia spectrum disorders: a cross-sectional study. Neuropsychiatr Dis Treat 2017;13:1201-1209 [FREE Full text] [doi: 10.2147/NDT.S130818] [Medline: 28490882]
18. Välimäki M, Kurki M, Hätönen H, Koivunen M, Selander M, Saarijärvi S, et al. Developing an internet-based support system for adolescents with depression. JMIR Res Protoc 2012 Dec 12;1(2):e22 [FREE Full text] [doi: 10.2196/resprot.2263] [Medline: 23612485]
19. Griffiths KM, Christensen H. Internet-based mental health programs: a powerful tool in the rural medical kit. Aust J Rural Health 2007 Apr;15(2):81-87. [doi: 10.1111/j.1440-1584.2007.00859.x] [Medline: 17441815]
20. Topolovec-Vranic J, Mansoor Y, Ennis N, Lightfoot D. Technology-adaptable interventions for treating depression in adults with cognitive impairments: protocol for a systematic review. Syst Rev 2015 Apr 7;4:42 [FREE Full text] [doi: 10.1186/s13643-015-0032-4] [Medline: 25875655]
21. Maideen SF, Mohd-Sidik S, Rampal L, Mukhtar F, Ibrahim N, Phang C, et al. A web-based psychoeducational intervention program for depression and anxiety in an adult community in Selangor, Malaysia: protocol of a randomized controlled trial. JMIR Res Protoc 2016 Jun 21;5(2):e112 [FREE Full text] [doi: 10.2196/resprot.4622] [Medline: 27329333]
22. Hidalgo-Mazzei D, Mateu A, Reinares M, Matic A, Vieta E, Colom F. Internet-based psychological interventions for bipolar disorder: review of the present and insights into the future. J Affect Disord 2015 Dec 1;188:1-13. [doi: 10.1016/j.jad.2015.08.005] [Medline: 26342885]
23. Rotondi AJ, Anderson CM, Haas GL, Eack SM, Spring MB, Ganguli R, et al. Web-based psychoeducational intervention for persons with schizophrenia and their supporters: one-year outcomes. Psychiatr Serv 2010 Nov;61(11):1099-1105. [doi: 10.1176/ps.2010.61.11.1099] [Medline: 21041348]
24. Steinwachs DM, Roter DL, Skinner EA, Lehman AF, Fahey M, Cullen B, et al. A web-based program to empower patients who have schizophrenia to discuss quality of care with mental health providers. Psychiatr Serv 2011 Nov;62(11):1296-1302 [FREE Full text] [doi: 10.1176/ps.62.11.pss6211_1296] [Medline: 22211208]
25. Glynn SM, Randolph ET, Garrick T, Lui A. A proof of concept trial of an online psychoeducational program for relatives of both veterans and civilians living with schizophrenia. Psychiatr Rehabil J 2010;33(4):278-287. [doi: 10.2975/33.4.2010.278.287] [Medline: 20374986]
26. Topolovec-Vranic J, Cullen N, Michalak A, Ouchterlony D, Bhalerao S, Masanic C, et al. Evaluation of an online cognitive behavioural therapy program by patients with traumatic brain injury and depression. Brain Inj 2010;24(5):762-772. [doi: 10.3109/02699051003709599] [Medline: 20370383]
27. Fraval A, Chandrananth J, Chong YM, Coventry LS, Tran P. Internet based patient education improves informed consent for elective orthopaedic surgery: a randomized controlled trial. BMC Musculoskelet Disord 2015 Feb 7;16:14 [FREE Full text] [doi: 10.1186/s12891-015-0466-9] [Medline: 25885962]
28. Athanasopoulou C, Hätönen H, Suni S, Lionis C, Griffiths KM, Välimäki M. An analysis of online health information on schizophrenia or related conditions: a cross-sectional survey. BMC Med Inform Decis Mak 2013 Aug 30;13:98 [FREE Full text] [doi: 10.1186/1472-6947-13-98] [Medline: 23992448]
29. Statistics Finland. 2017. Mobile Phone is Becoming Increasingly Popular for Using the Internet - Use Purposes Are Diversifying URL: http://www.stat.fi/til/sutivi/2017/13/sutivi_2017_13_2017-11-22_tie_001_en.html [accessed 2019-08-30]
30. Kannisto KA, Korhonen J, Adams CE, Koivunen MH, Vahlberg T, Välimäki MA. Factors associated with dropout during recruitment and follow-up periods of a mhealth-based randomized controlled trial for mobile.net to encourage treatment adherence for people with serious mental health problems. J Med Internet Res 2017 Feb 21;19(2):e46 [FREE Full text] [doi: 10.2196/jmir.6417] [Medline: 28223262]
31. Berry N, Lobban F, Emsley R, Bucci S. Acceptability of interventions delivered online and through mobile phones for people who experience severe mental health problems: a systematic review. J Med Internet Res 2016 May 31;18(5):e121 [FREE Full text] [doi: 10.2196/jmir.5250] [Medline: 27245693]

32. Anttila M, Välimäki M, Hätönen H, Luukkaala T, Kaila M. Use of web-based patient education sessions on psychiatric wards. Int J Med Inform 2012 Jun;81(6):424-433. [doi: 10.1016/j.ijmedinf.2012.02.004] [Medline: 22381805]
33. Ben-Zeev D, Brian RM, Jonathan G, Razzano L, Pashka N, Carpenter-Song E, et al. Mobile health (mhealth) versus clinic-based group intervention for people with serious mental illness: a randomized controlled trial. Psychiatr Serv 2018 Sep 1;69(9):978-985. [doi: 10.1176/appi.ps.201800063] [Medline: 29793397]
34. Schlosser DA, Campellone TR, Truong B, Etter K, Vergani S, Komaiko K, et al. Efficacy of PRIME, a mobile app intervention designed to improve motivation in young people with schizophrenia. Schizophr Bull 2018 Aug 20;44(5):1010-1020 [FREE Full text] [doi: 10.1093/schbul/sby078] [Medline: 29939367]
35. Bucci S, Barrowclough C, Ainsworth J, Machin M, Morris R, Berry K, et al. Actissist: proof-of-concept trial of a theory-driven digital intervention for psychosis. Schizophr Bull 2018 Aug 20;44(5):1070-1080 [FREE Full text] [doi: 10.1093/schbul/sby032] [Medline: 29566206]
36. Killikelly C, He Z, Reeder C, Wykes T. Improving adherence to web-based and mobile technologies for people with psychosis: systematic review of new potential predictors of adherence. JMIR Mhealth Uhealth 2017 Jul 20;5(7):e94 [FREE Full text] [doi: 10.2196/mhealth.7088] [Medline: 28729235]
37. Kanuch SW, Cassidy KA, Dawson NV, Athey M, Fuentes-Casiano E, Sajatovic M. Recruiting and retaining individuals with serious mental illness and diabetes in clinical research: lessons learned from a randomized, controlled trial. J Health Dispar Res Pract 2016;9(3):115-126 [FREE Full text] [Medline: 28533944]
38. Terp M, Jørgensen R, Laursen BS, Mainz J, Bjørnes CD. A smartphone app to foster power in the everyday management of living with schizophrenia: qualitative analysis of young adults' perspectives. JMIR Ment Health 2018 Oct 1;5(4):e10157 [FREE Full text] [doi: 10.2196/10157] [Medline: 30274966]
39. Abbott M, Arthur A, Walker L, Doody G. The challenge of recruiting people with schizophrenia to a health promotion trial. Br J Gen Pract 2005 Aug;55(517):634-636 [FREE Full text] [Medline: 16105374]
40. Lee H, Schepp KG. Lessons learned from research with adolescents with schizophrenia and their families. Arch Psychiatr Nurs 2013 Aug;27(4):198-203. [doi: 10.1016/j.apnu.2013.03.002] [Medline: 23915698]
41. Monsen KA. Intervention Effectiveness Research: Quality Improvement and Program Evaluation. First Edition. Cham, Germany: Springer; 2018.
42. World Health Organization. 2016. Monitoring and Evaluating Digital Health Interventions - A Practical Guide to Conducting Research and Assessment URL: https://apps.who.int/iris/bitstream/handle/10665/252183/9789241511766-eng.pdf;jsessionid=7F11859CFC39E6A70B52EC49F000B75E?sequence=1 [accessed 2019-08-30]
43. Jadad AR, Moore RA, Carroll D, Jenkinson C, Reynolds DJ, Gavaghan DJ, et al. Assessing the quality of reports of randomized clinical trials: is blinding necessary? Control Clin Trials 1996 Feb;17(1):1-12. [doi: 10.1016/0197-2456(95)00134-4] [Medline: 8721797]
44. Keogh-Brown MR, Bachmann MO, Shepstone L, Hewitt C, Howe A, Ramsay CR, et al. Contamination in trials of educational interventions. Health Technol Assess 2007 Oct;11(43):iii, ix-iii,107 [FREE Full text] [doi: 10.3310/hta11430] [Medline: 17935683]
45. Edwards SJ, Braunholtz DA, Lilford RJ, Stevens AJ. Ethical issues in the design and conduct of cluster randomised controlled trials. Br Med J 1999 May 22;318(7195):1407-1409 [FREE Full text] [doi: 10.1136/bmj.318.7195.1407] [Medline: 10334756]
46. Harris AD, McGregor JC, Perencevich EN, Furuno JP, Zhu J, Peterson DE, et al. The use and interpretation of quasi-experimental studies in medical informatics. J Am Med Inform Assoc 2006;13(1):16-23 [FREE Full text] [doi: 10.1197/jamia.M1749] [Medline: 16221933]
47. World Health Organization. 2016. nternational Statistical Classification of Diseases and Related Health Problems. Tenth Revision URL: https://icd.who.int/browse10/2016/en [accessed 2019-08-30]
48. The Council for International Organizations of Medical Sciences (CIOMS). 2016. International Ethical Guidelines for Health-Related Research Involving Humans URL: https://cioms.ch/wp-content/uploads/2017/01/WEB-CIOMS-EthicalGuidelines.pdf [accessed 2019-08-30]
49. World Medical Association. World Medical Association declaration of Helsinki: ethical principles for medical research involving human subjects. J Am Med Assoc 2013 Nov 27;310(20):2191-2194. [doi: 10.1001/jama.2013.281053] [Medline: 24141714]
50. Advisory Board on Research Ethics: TENK. 2012. Responsible Conduct of Research and Procedures for Handling Allegations of Misconduct in Finland URL: http://www.tenk.fi/sites/tenk.fi/files/HTK_ohje_2012.pdf [accessed 2019-08-30]
51. Etusivu - FINLEX ®. Medical Research Act 488/1999 URL: https://www.finlex.fi/fi/laki/kaannokset/1999/en19990488.pdf [accessed 2019-08-30]
52. Etusivu - FINLEX ®. Personal Data Act 523/1999 URL: http://finlex.fi/en/laki/kaannokset/1999/en19990523_20000986.pdf [accessed 2019-08-30]
53. Välimäki M, Anttila M, Hätönen H, Koivunen M, Jakobsson T, Pitkänen A, et al. Design and development process of patient-centered computer-based support system for patients with schizophrenia spectrum psychosis. Inform Health Soc Care 2008 Jun;33(2):113-123. [doi: 10.1080/17538150802127207] [Medline: 18604755]

54. Laine A, Anttila M, Välimäki M. Modification of an internet-based patient education program for adults with schizophrenia spectrum disorder to suit adolescents with psychosis. Inform Health Soc Care 2016;41(3):230-246. [doi: 10.3109/17538157.2015.1008487] [Medline: 25710517]
55. Health On the Net. [Mind.Net] Mieli.Net - [Support System for Mental Health Work] Tukijärjestelmä mielenterveystyöhön URL: https://www.healthonnet.org/HONcode/Finnish/?HONConduct193826 [accessed 2019-09-03]
56. Laine A, Välimäki M, Löyttyniemi E, Pekurinen V, Marttunen M, Anttila M. The impact of a web-based course concerning patient education for mental health care professionals: quasi-experimental study. J Med Internet Res 2019 Mar 1;21(3):e11198 [FREE Full text] [doi: 10.2196/11198] [Medline: 30821697]
57. Stiles WB, Gordon LE, Lani JA. Session evaluation and the session evaluation questionnaire. In: Tryon GS, editor. Counseling Based On Process Research: Applying What We Know. Boston, Massachusetts: Allyn & Bacon; 2002:325-343.
58. Reynolds Jr DJ, Stiles WB, Grohol JM. An investigation of session impact and alliance in internet based psychotherapy: preliminary results. Counsell Psychother Res 2006 Sep;6(3):164-168. [doi: 10.1080/14733140600853617]
59. Sperber AD. Translation and validation of study instruments for cross-cultural research. Gastroenterology 2004 Jan;126(1 Suppl 1):S124-S128. [doi: 10.1053/j.gastro.2003.10.016] [Medline: 14978648]
60. Hilton A, Skrutkowski M. Translating instruments into other languages: development and testing processes. Cancer Nurs 2002 Feb;25(1):1-7. [doi: 10.1097/00002820-200202000-00001] [Medline: 11838715]
61. Goldberg L, Lide B, Lowry S, Massett HA, O'Connell T, Preece J, et al. Usability and accessibility in consumer health informatics current trends and future challenges. Am J Prev Med 2011 May;40(5 Suppl 2):S187-S197. [doi: 10.1016/j.amepre.2011.01.009] [Medline: 21521594]
62. US Department of Health & Human Services. Research-Based Web Design & Usability Guidelines. Washington, DC: US Government Printing Office; 2006.
63. Schwarzer R, Jerusalem M. Generalized self-efficacy scale. In: Johnston M, Wright SC, Weinman J, editors. Measures in Health Psychology: A User's Portfolio. Windsor, England: Nfer-Nelson; 1995:35-37.
64. Scholz U, Doña BG, Sud S, Schwarzer R. Is general self-efficacy a universal construct? Eur J Psychol Assess 2002 Sep 1;18(3):242-251. [doi: 10.1027//1015-5759.18.3.242]
65. Rosenberg M. Society and the Adolescent Self-Image. Princeton, NJ: Princeton University Press; 1965.
66. Schmitt DP, Allik J. Simultaneous administration of the Rosenberg self-esteem scale in 53 nations: exploring the universal and culture-specific features of global self-esteem. J Pers Soc Psychol 2005 Oct;89(4):623-642. [doi: 10.1037/0022-3514.89.4.623] [Medline: 16287423]
67. Evers AW, Kraaimaat FW, van Lankveld W, Jongen PJ, Jacobs JW, Bijlsma JW. Beyond unfavorable thinking: the illness cognition questionnaire for chronic diseases. J Consult Clin Psychol 2001 Dec;69(6):1026-1036. [doi: 10.1037/0022-006X.69.6.1026] [Medline: 11777106]
68. Lauwerier E, Crombez G, van Damme S, Goubert L, Vogelaers D, Evers AW. The construct validity of the illness cognition questionnaire: the robustness of the three-factor structure across patients with chronic pain and chronic fatigue. Int J Behav Med 2010 Jun;17(2):90-96. [doi: 10.1007/s12529-009-9059-z] [Medline: 19757084]
69. Ascher-Svanum H. Development and validation of a measure of patients' knowledge about schizophrenia. Psychiatr Serv 1999 Apr;50(4):561-563. [doi: 10.1176/ps.50.4.561] [Medline: 10211743]
70. Etusivu - FINLEX ®. Mental Health Act 1116/1990 URL: https://www.finlex.fi/en/laki/kaannokset/1990/en19901116 [accessed 2019-08-30]
71. Choi NG, Dinitto DM. The digital divide among low-income homebound older adults: internet use patterns, ehealth literacy, and attitudes toward computer/internet use. J Med Internet Res 2013 May 2;15(5):e93 [FREE Full text] [doi: 10.2196/jmir.2645] [Medline: 23639979]
72. Cohen J. Statistical Power Analysis for the Behavioral Sciences. Second Edition. New York: Routledge; 1988.
73. Easy Calculation. Effect Size Calculator URL: https://www.easycalculation.com/statistics/effect-size.php [accessed 2019-08-30]
74. Kaminsky A, Roberts LW, Brody JL. Influences upon willingness to participate in schizophrenia research: an analysis of narrative data from 63 people with schizophrenia. Ethics Behav 2003;13(3):279-302. [doi: 10.1207/S15327019EB1303_06] [Medline: 14680009]
75. Edlinger M, Deisenhammer EA, Fiala M, Hofer A, Kemmler G, Strauss R, et al. Attitudes of patients with schizophrenia and depression towards psychiatric research. Psychiatry Res 2010 May 15;177(1-2):172-176. [doi: 10.1016/j.psychres.2008.12.010] [Medline: 20362343]
76. Schäfer I, Burns T, Fleischhacker WW, Galderisi S, Rybakowski JK, Libiger J, et al. Attitudes of patients with schizophrenia and depression to psychiatric research: a study in seven European countries. Soc Psychiatry Psychiatr Epidemiol 2011 Feb;46(2):159-165. [doi: 10.1007/s00127-010-0181-7] [Medline: 20119828]
77. Woodall A, Morgan C, Sloan C, Howard L. Barriers to participation in mental health research: are there specific gender, ethnicity and age related barriers? BMC Psychiatry 2010 Dec 2;10:103 [FREE Full text] [doi: 10.1186/1471-244X-10-103] [Medline: 21126334]
78. Loue S, Sajatovic M. Research with severely mentally ill Latinas: successful recruitment and retention strategies. J Immigr Minor Health 2008 Apr;10(2):145-153. [doi: 10.1007/s10903-007-9063-9] [Medline: 17636454]

79. Miller BJ, Stewart A, Schrimsher J, Peeples D, Buckley PF. How connected are people with schizophrenia? Cell phone, computer, email, and social media use. Psychiatry Res 2015 Feb 28;225(3):458-463. [doi: 10.1016/j.psychres.2014.11.067] [Medline: 25563669]
80. American Psychiatric Association. What Is Schizophrenia? URL: https://www.psychiatry.org/patients-families/schizophrenia/what-is-schizophrenia [accessed 2019-08-30]
81. Furimsky I, Cheung AH, Dewa CS, Zipursky RB. Strategies to enhance patient recruitment and retention in research involving patients with a first episode of mental illness. Contemp Clin Trials 2008 Nov;29(6):862-866. [doi: 10.1016/j.cct.2008.07.005] [Medline: 18721902]
82. Jørgensen R, Munk-Jørgensen P, Lysaker PH, Buck KD, Hansson L, Zoffmann V. Overcoming recruitment barriers revealed high readiness to participate and low dropout rate among people with schizophrenia in a randomized controlled trial testing the effect of a guided self-determination intervention. BMC Psychiatry 2014 Feb 3;14:28 [FREE Full text] [doi: 10.1186/1471-244X-14-28] [Medline: 24490977]
83. Kivlighan DM, Marmarosh CL, Hilsenroth MJ. Client and therapist therapeutic alliance, session evaluation, and client reliable change: a moderated actor-partner interdependence model. J Couns Psychol 2014 Jan;61(1):15-23. [doi: 10.1037/a0034939] [Medline: 24188649]
84. Reynolds Jr DJ, Stiles WB, Bailer AJ, Hughes MR. Impact of exchanges and client-therapist alliance in online-text psychotherapy. Cyberpsychol Behav Soc Netw 2013 May;16(5):370-377 [FREE Full text] [doi: 10.1089/cyber.2012.0195] [Medline: 23530546]
85. Bavojdan MR, Towhidi A, Rahmati A. The relationship between mental health and general self-efficacy beliefs, coping strategies and locus of control in male drug abusers. Addict Health 2011;3(3-4):111-118 [FREE Full text] [doi: 10.22122/ahj.v3i3-4.69] [Medline: 24494125]
86. Kershaw T, Ellis KR, Yoon H, Schafenacker A, Katapodi M, Northouse L. The interdependence of advanced cancer patients' and their family caregivers' mental health, physical health, and self-efficacy over time. Ann Behav Med 2015 Dec;49(6):901-911 [FREE Full text] [doi: 10.1007/s12160-015-9743-y] [Medline: 26489843]
87. Anttila M, Koivunen M, Välimäki M. Information technology-based standardized patient education in psychiatric inpatient care. J Adv Nurs 2008 Oct;64(2):147-156. [doi: 10.1111/j.1365-2648.2008.04770.x] [Medline: 18990095]
88. Koivunen M, Saranto K. Nursing professionals' experiences of the facilitators and barriers to the use of telehealth applications: a systematic review of qualitative studies. Scand J Caring Sci 2018 Mar;32(1):24-44. [doi: 10.1111/scs.12445] [Medline: 28771752]
89. National Institute for Health and Welfare. 2018. Psychiatric Specialist Medical Care 2017 URL: http://www.julkari.fi/bitstream/handle/10024/136978/tr33_18.pdf?sequence=5&isAllowed=y [accessed 2019-08-30]
90. Hamric AB, Hanson CM, Tracy MF, O'Grady ET. Advanced Practice Nursing: An Integrative Approach. St Louis, Missouri: Elsevier Saunders; 2014.
91. Hartmann A, Joos A, Orlinsky DE, Zeeck A. Accuracy of therapist perceptions of patients' alliance: exploring the divergence. Psychother Res 2015;25(4):408-419. [doi: 10.1080/10503307.2014.927601] [Medline: 25000227]

Abbreviations

GSE: General Self-Efficacy Scale
ICD-10: International Classification of Diseases, 10th Revision
SEQ: Session Evaluation Questionnaire (Form 5)
SES: Rosenberg Self-Esteem Scale

Edited by G Eysenbach; submitted 10.12.18; peer-reviewed by A Rotondi, M Katapodi, T Campellone; comments to author 08.04.19; revised version received 25.06.19; accepted 19.08.19; published 17.10.19.

Please cite as:
Laine A, Välimäki M, Pekurinen V, Löyttyniemi E, Marttunen M, Anttila M
Feasibility, Acceptability, and Preliminary Impacts of Web-Based Patient Education on Patients With Schizophrenia Spectrum Disorder: Quasi-Experimental Cluster Study
J Med Internet Res 2019;21(10):e13073
URL: http://www.jmir.org/2019/10/e13073/
doi:10.2196/13073
PMID:31625952

Original Paper

A Comparison of Users and Nonusers of a Web-Based Intervention for Carers of Older Persons With Alzheimer Disease and Related Dementias: Mixed Methods Secondary Analysis

Wendy Duggleby[1], PhD; Jenny Ploeg[2], PhD; Carrie McAiney[3], PhD; Kathryn Fisher[2], PhD; Kathya Jovel Ruiz[1], BSc; Sunita Ghosh[4], PhD; Shelley Peacock[5], PhD; Maureen Markle-Reid[2], PhD; Allison Williams[6], PhD; Jean Triscott[7], MD; Jennifer Swindle[1], PhD

[1]Faculty of Nursing, University of Alberta, Edmonton, AB, Canada
[2]School of Nursing, Faculty of Health Sciences, McMaster University, Hamilton, ON, Canada
[3]School of Public Health and Health Systems, University of Waterloo, Waterloo, ON, Canada
[4]Cancer Care, Alberta Health Services, Edmonton, AB, Canada
[5]College of Nursing, University of Saskatchewan, Saskatoon, SK, Canada
[6]School of Geography and Earth Sciences, McMaster University, Hamilton, ON, Canada
[7]Faculty of Medicine and Dentistry, University of Alberta, Edmonton, AB, Canada

Corresponding Author:
Wendy Duggleby, PhD
Faculty of Nursing
University of Alberta
11405 87th Ave
Edmonton, AB, T6G 1C9
Canada
Phone: 1 7804928660
Email: wendy.duggleby@ualberta.ca

Abstract

Background: A self-administered Web-based intervention was developed to help carers of persons with Alzheimer disease and related dementias (ADRD) and multiple chronic conditions (MCC) deal with the significant transitions they experience. The intervention, My Tools 4 Care (MT4C), was evaluated during a pragmatic mixed methods randomized controlled trial with 199 carers. Those in the intervention group received free, password-protected access to MT4C for three months. MT4C was found to increase hope in participants at three months compared with the control group. However, in the intervention group, 22% (20/92) did not use MT4C at all during the three-month period.

Objective: This mixed methods secondary analysis aimed to (1) examine differences at three months in the outcomes of hope, self-efficacy, and health-related quality of life (HRQOL) scores in users (ie, those who used MT4C at least once during the three-month period) compared with nonusers and (2) identify reasons for nonuse.

Methods: Data from the treatment group of a pragmatic mixed methods randomized controlled trial were used. Through audiotaped telephone interviews, trained research assistants collected data on participants' hope (Herth Hope Index; HHI), self-efficacy (General Self-Efficacy Scale; GSES), and HRQOL (Short-Form 12-item health survey version 2; SF-12v2) at baseline, one month, and three months. Treatment group participants also provided feedback on MT4C through qualitative telephone interviews at one month and three months. Analysis of covariance was used to determine differences at three months, and generalized estimating equations were used to determine significant differences in HHI, GSES, and SF-12v2 between users and nonusers of MT4C from baseline to three months. Interview data were analyzed using content analysis and integrated with quantitative data at the result stage.

Results: Of the 101 participants at baseline, 9 (9%) withdrew from the study, leaving 92 participants at three months of which 72 (78%) used MT4C at least once; 20 (22%) participants did not use it at all. At baseline, there were no statistically significant differences in demographic characteristics and in outcome variables (HHI, GSES, and SF-12v2 mental component score and physical component score) between users and nonusers. At three months, participants who used MT4C at least once during the three-month period (users) reported higher mean GSES scores (P=.003) than nonusers. Over time, users had significantly higher

GSES scores than nonusers (*P*=.048). Reasons for nonuse of MT4C included the following: caregiving demands, problems accessing MT4C (poor connectivity, computer literacy, and navigation of MT4C), and preferences (for paper format or face-to-face interaction).

Conclusions: Web-based interventions, such as MT4C, have the potential to increase the self-efficacy of carers of persons with ADRD and MCC. Future research with MT4C should consider including educational programs for computer literacy and providing alternate ways to access MT4C in addition to Web-based access.

Trial Registration: ClinicalTrials.gov NCT02428387; https://clinicaltrials.gov/ct2/show/NCT02428387

(J Med Internet Res 2019;21(10):e14254) doi:10.2196/14254

KEYWORDS

Web-based intervention; carers; dementia; multiple chronic conditions; program evaluation

Introduction

Family carers (unpaid family or friends) of persons with Alzheimer disease and related dementias (ADRD) have been recognized worldwide as providers of the majority of care [1]. The need to support these family carers is well documented as they experience significant changes in their lives [2] that can negatively impact their physical and mental health [3,4]. Family carers have been found to seek information for themselves and others using computers, smartphones, or other electronic means more frequently than noncarers [5]. Web-based interventions to support family carers are increasingly becoming available and affordable, flexible, and accessible [6]. In a recent meta-analysis, Web-based interventions for family carers were found to have positive outcomes such as increased mental health [7] and self-efficacy [8]. However, carers have also reported barriers to using Web-based interventions such as difficulty with language and computer literacy [9] and with navigation of the websites [10]. The limited interaction with other carers when using Web-based interventions has also been a concern [11,12].

A Web-based intervention, My Tools 4 Care (MT4C), was developed to support family carers of persons with ADRD and multiple chronic conditions (MCC). MT4C was initially developed as a hard copy workbook that was piloted and showed promise in terms of helping family carers of persons with dementia [13]. The next step was to work with Atmist (Web developers), the research team, and family carers to develop a feasible, acceptable, and easy-to-use Web-based version. MT4C [14] is a self-administered, flexible, tailored intervention as carers decide which activities they wish to engage in and when. It can be used on a computer, tablet, or smartphone. MT4C was evaluated during a pragmatic mixed methods randomized control trial with 199 carers between June 2015 and April 2017 and was found to significantly increase participants' hope in the treatment group compared with the control group [15].

Of 92 participants in the treatment group, 20 (22%) did not use MT4C over the three-month period. This is similar to other Web-based internet intervention studies that reported a substantial number of treatment group participants who did not use the interventions [10,12]. The Consolidated Standards of Reporting Trials-Electronic Health guidelines for reporting Web-based intervention trials recommend the common practice of using an intent-to-treat analysis in trials that include users and nonusers [16]. However, it is also important to examine the data of nonuser participants [17]. This paper reports a secondary analysis to provide insight into the characteristics and the difference in outcomes of hope, general self-efficacy, and quality of life in users versus nonusers of MT4C. This examination of MT4C will inform the evaluation of future Web-based interventions for family carers of persons with ADRD and MCC.

The aim of this secondary analysis was to examine differences in outcomes (hope, self-efficacy, and quality of life) in participants who used MT4C (users) in a three-month period versus those who did not use it (nonusers) and to examine reasons for nonuse. The following research questions guided the study:

1. Was there a significant difference in demographic characteristics in the users versus nonusers?
2. Was there a significant difference in hope, self-efficacy, and quality of life at three months in users versus nonusers?
3. What were the reasons for nonuse from the qualitative data collected for nonusers?

Those who did not use MT4C were not a control group, as they had the opportunity to use it. The nonusers had been randomly assigned to the treatment group but chose not to use MT4C. We hypothesized that the users of MT4C will report a statistically significant increase in hope, self-efficacy, and quality of life compared with the nonusers.

Methods

Design

A detailed protocol [18] and 2 articles describing MT4C and its evaluation have been published elsewhere [15,19]; thus, only details relevant to this secondary analysis are provided herein. Similar to the pragmatic trial, this study utilized a mixed methods comparative design. The data reported here focus on family carers allocated to the treatment group (N=101). Qualitative data from the interviews and quantitative data from the treatment group collected during the original study were integrated in the results stage. Qualitative data were used to understand the quantitative results.

Ethics

The primary study received ethical approval from the University of Alberta Health Research Ethics Board (No. Pro00048721) and the Hamilton Integrated Research Ethics Board (No.

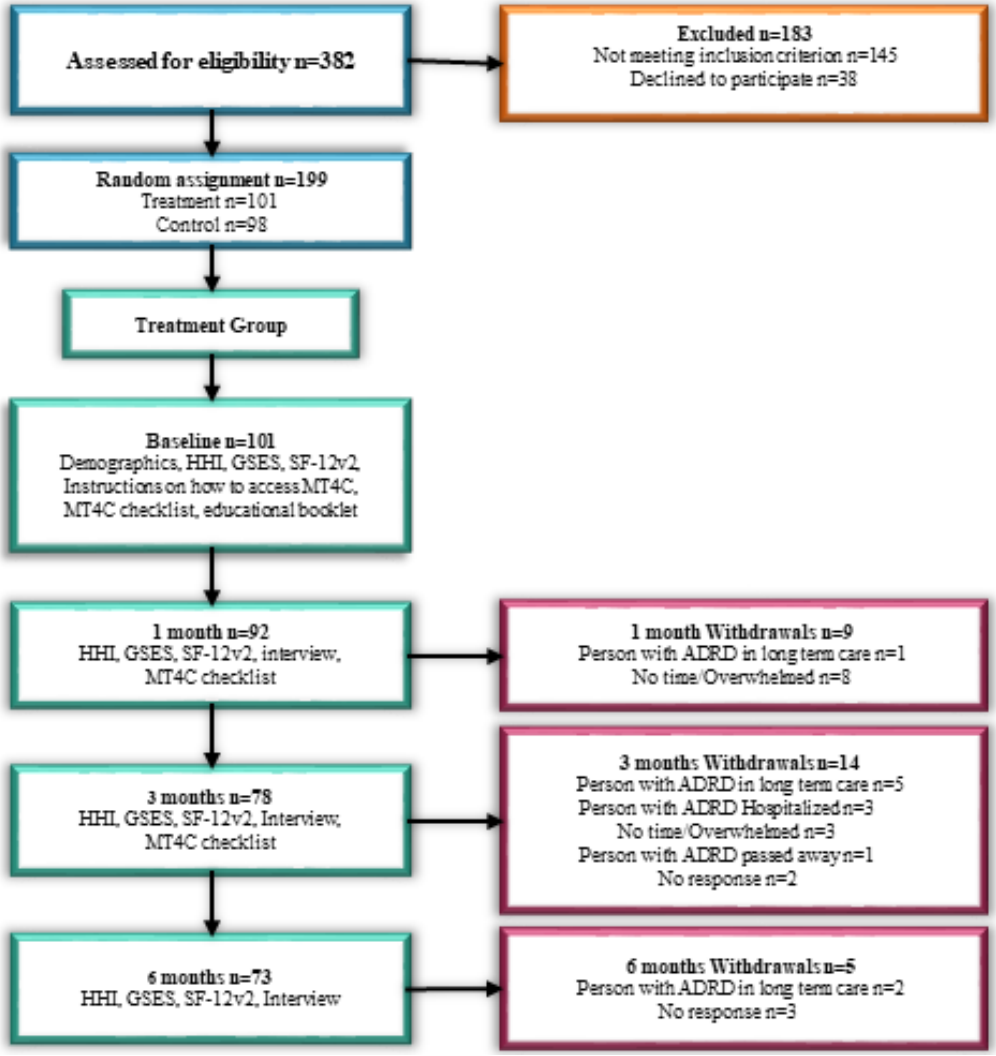
15-309). The initial ethics application included the ability to conduct a secondary analysis of the data.

Recruitment of Participants

Family carers were invited to participate if they were over the age of 18 years and were caring for a person aged 65 years or older living with ADRD and MCC in the community. In addition, they needed to have a valid email address and access to a computer. Family carers were recruited through multiple community organizations including Alzheimer Society branches in each province and advertisements in local community newspapers in Alberta. If they met the eligibility criteria, they were asked to contact the researchers. Participants were randomly assigned to a treatment or control group using stratified permuted block randomization. Different consent forms (one for the treatment group and one for the control group) were used to blind participants regarding their group assignment.

Once consent was obtained, trained data collectors collected all data (baseline, one month, and three months) via telephone and entered it into REDCap, a secure, password-protected Web-based data collection service, offered at the University of Alberta. Data collection occurred from June 2015 to April 2017 and is reported in more detail in the study protocol [18]. Data collection procedures for the treatment group and the number of participants at each period are presented in Figure 1.

Figure 1. Data collection procedures and numbers of participants. ADRD: Alzheimer disease and related dementias; GSES: general self-efficacy scale; HHI: herth hope index; MT4C: My Tools 4 Care; SF-12v2: short-form 12-item health survey.



Intervention

Following baseline interviews, participants in the treatment group received free, password-protected access to MT4C for three months. Research assistants, using a standardized script, instructed the participants to access MT4C at their convenience on a computer, tablet, or smartphone. A follow-up email was sent to participants in the treatment group with instructions on how to access the site and their login information. After logging in, the first page provided instructions on how to use MT4C. Each Web page also contained a menu outlining the sections that comprised MT4C: (1) about me; (2) common changes to expect; (3) frequently asked questions; (4) resources; (5) important health information; and (6) calendar. MT4C also provided options to add formatted text, photos, and PDF files in certain sections. All information entered by participants was treated confidentially and was not accessible to the study team. Participants also received an electronic copy of the Alzheimer Society's *The Progression of Alzheimer's Disease* booklet [20], a copy of the study questionnaires, and the MT4C toolkit

checklist intended for participants to record their use of the MT4C site. During the one-month interview, trained data collectors encouraged nonusers to use MT4C. No changes or alterations were made to MT4C during the study.

Measures

Data Collection

Data collected at baseline included age, gender, years in caregiver role, employment status, ethnicity, household income before taxes, living arrangement, the relationship to the person with ADRD and MCC, and any assistance with caregiving. Data regarding sex, age, and number of chronic conditions of the person with ADRD and MCC were also collected.

Data on family carers' hope, self-efficacy, and health-related quality of life (HRQOL) were collected at baseline, one month, three months, and 6 months by trained research assistants using the measures outlined in the following section. Figure 1 outlines the data collection procedures.

Herth Hope Index

To measure hope using the Herth Hope Index (HHI), participants answered 12 questions scored on a Likert-type scale from 1 (strongly disagree) to 4 (strongly agree). A total hope score (range 12-48) was reported (higher scores indicate higher hope) along with 3 subscales: (1) temporality and future, (2) positive readiness and expectancy, and (3) interconnectedness. The HHI has a test-retest reliability of 0.91 (P<.05) and criterion-related validity r of 0.81 to 0.92 (P<.005) [21].

General Self-Efficacy Scale

The General Self-Efficacy Scale (GSES) is a measure of perceived self-efficacy or belief that one can deal with difficult tasks or cope with adversity using a 10-item 4-point scale [22]. Total scores ranged from 10 to 40. It is a reliable tool with a Cronbach alpha coefficient ranging from .76 to .90 (P<.05).

Short Form 12-Item Health Survey

The SF-12v2 is a measure of HRQOL, consisting of 12 questions measuring 8 domains of well-being and functioning (physical functioning, role functioning, bodily pain, general health, vitality, social functioning, emotional health, and mental health) [23,24]. Responses to the 12 questions are summarized by 2 scores: physical component summary score (PCS; estimated test-retest reliability of r=0.89) and a mental component summary score (MCS; estimated test-retest reliability r=0.86) that range from 0 to 100 [25].

Qualitative Interviews

Interviews were semistructured and completed over the telephone by a trained research assistant. All interviews were audio-recorded and transcribed verbatim by an experienced transcriptionist. Participants were asked questions such as *What were you thinking about when you worked on MT4C?*; *Did it help you deal with significant changes?*; *What did you like best?*; and *What did you like least?* As indicated by the larger study protocol, qualitative interviews were conducted with a subsample of study participants. For those in the treatment group, 6 of the 20 nonusers were interviewed using semistructured interviews.

My Tools 4 Care Checklist (Use of My Tools 4 Care)

The MT4C checklist was developed by the research team and was used to collect data on the participants' use of MT4C. The checklist was intended for participants to keep track of the number of times they accessed each section of MT4C and the amount of time spent on each section. Data from the checklist were used to determine the use of MT4C at one month and three months.

Participants also made comments on the checklist about their nonuse, which were considered qualitative data for this study.

Data Analysis

Data were entered in SPSS version 24 (IBM) and checked for accuracy by a trained research assistant. Before data analysis, participants were divided into 2 groups: (1) participants who used MT4C at least once over three months and (2) participants who did not use MT4C within the three-month period. Use of MT4C was captured using a dichotomous variable, where 1=used MT4C at least once during the three-month intervention period and 0=did not use MT4C during the three-month intervention period.

Participant Characteristics

Means and SDs were used to represent continuous demographic characteristics of participants and persons with ADRD and MCC; categorical data were reported with numbers and percentages. Chi-squared statistical analysis and t tests were used to determine differences in demographic characteristics between the groups.

Outcome Measures

Analysis of covariance (ANCOVA) was used to test the differences in outcome variables between users and nonusers at three months. Separate ANCOVA models were run for each outcome, with the three-month outcome as the dependent variable, group (users and nonusers) as the independent variable, and baseline value of the outcome as the covariate. A P value of <.05 was used for statistical significance, and 2-sided tests were used. A complete case analysis was used, which means we did not impute for missing data (ie, we used people who had a complete record for the 3 time points baseline, one month, and three months).

Generalized estimating equations (GEE) were used to determine differences between the 2 groups (users vs nonusers) over time for the main outcome variables of HHI, GSES, and SF-12v2 MCS and PCS. GEE is an alternative statistical method appropriate for repeated measures data and is more flexible than other methods (eg, repeated measures ANCOVA) because it does not require that outcomes be normally distributed and can handle both continuous and dichotomous outcomes. It can also be used with small sample sizes [26]. Use was captured dichotomously at 3 time points in the GEE models, with use=0 for all intervention group participants at baseline, and 1 (used) or 0 (did not use) at one month and three months (depending on use reported by participants at these time points). Separate GEE models were run for each outcome (primary and secondary).

Reasons for Nonuse of My Tools 4 Care

Nonusers' MT4C checklist and qualitative data from interviews were analyzed using content analysis [27] and informed the quantitative data in the results phase. Transcripts were read overall by a trained research assistant and organized into categories to address the study purpose. Trustworthiness of the data was maintained by keeping an audit trail and using participants' words as much as possible.

Results

Comparison of User and Nonuser Participants

A total of 101 participants were allocated to the treatment group at baseline. Following baseline measures, 9 participants withdrew and the remaining 92 participants received instructions on how to access MT4C. Figure 1 illustrates the number of persons at baseline and three months. The mean age of all participants in the treatment group was 63.5 years (SD 12.0), and they had been carers for an average of 4.1 years (SD 3.9). The majority of participants were female (73/92, 79%), white (84/92, 91%), were living with a person with ADRD (63/92, 62%), and were the spouse of a person with ADRD (48/92, 52%). No statistically significant differences were found in the demographic characteristics of users and nonusers (Table 1).

The means and SDs of the outcome measures (HHI, GSES, MCS, and PCS for each group at baseline and three months) are presented in Table 2.

Table 3 provides the ANCOVA results for each outcome. The group variable (users and nonusers) was significant for the GSES outcome (P=.003), indicating that the use of MT4C during the three-month period was associated with an increase in GSES from baseline to three months. The use of MT4C was not associated with significant differences in the other outcomes.

Table 4 provides the GEE model results and shows that the use of MT4C was associated with an increase in GSES over three months (P=.048). The use of MT4C was not associated with significant changes in the other outcomes.

Table 1. My Tools 4 Care users versus nonusers: baseline comparison characteristics.

Characteristics	Used MT4C[a] (N=72)	Did not use MT4C (N=20)	Total sample (N=92)	*P* value
Carers				
Gender, n (%)				
Male	14 (19)	5 (25)	19 (20)	.59
Female	58 (80)	15 (75)	73 (79)	.59
Age (years), mean (SD)	62.8 (12.2)	65.8 (11.3)	63.5 (12.0)	.38
Caregiving (years), mean (SD)	3.9 (4.0)	4.8 (3.6)	4.1 (3.9)	.38
Education (years), mean (SD)	14.2 (2.9)	14.2 (2.8)	14.2 (2.9)	.10
Chronic conditions, mean (SD)	2.3 (1.6)	2.2 (1.3)	2.2 (1.6)	.86
Marital status, n (%)				
Married or living with someone	60 (83)	17 (85)	77 (84)	.87
Single, widowed, divorced/separated	12 (17)	3 (15)	15 (16)	.87
Ethnicity, n (%)				
White	66 (9)	18 (90)	84 (9)	.97
Other	5 (7)	2 (10)	7 (8)	.97
Employed, n (%)				
Yes	31 (4)	3 (15)	34 (3)	.05
No	41 (57)	16 (84)	57 (6)	.05
Living with care recipient, n (%)				
Yes	50 (6)	13 (6)	63 (68)	.71
No	22 (31)	7(35)	29 (31)	.71
Relationship to care recipient, n (%)				
Husband/wife/life partner	37 (51)	11 (55)	48 (52)	.77
Other	35 (9)	9 (45.0)	44 (48)	.77
Finances meet needs, n (%)				
Completely, very well, adequately	60 (8)	14 (7)	74 (80)	.18
With some difficulty, not very well, totally inadequate	12 (17)	6 (30)	18 (19)	.18
Household income, n (%)				
<40,000	17 (24)	5 (25)	22 (24)	.99
>40,000 and <70,000	18 (25)	4 (20)	22 (24)	.99
>70,000	26 (36)	8 (40)	34 (37)	.99
No response	11 (15)	3 (15)	14 (15)	.99
Assistance with caring, n (%)				
Yes	49 (68)	15 (75)	64 (70)	.55
No	23 (32)	5 (25)	28 (30)	.55
Care recipient				
Gender, mean (%)				
Male	37 (51)	11 (55)	48 (52)	.77
Female	35 (49)	9 (45)	44 (48)	.77
Age (years), mean (SD)	79.6 (7.7)	82.5 (6.5)	80.2 (7.5)	.13
Chronic conditions, mean (SD)	10.6 (4.2)	10.1 (4.1)	10.5 (4.2)	.63

[a]MT4C: My Tools 4 Care.

Table 2. Mean and SD of outcomes at baseline and three months for users and nonusers.

Outcomes	Users MT4C[a] (N=72), mean (SD)	Nonusers MT4C (N=20), mean (SD)
Outcomes at baseline		
PCS[b]	50.80 (12.01)	50.51 (9.62)
MCS[c]	46.41(10.38)	44.19 (11.66)
HHI[d]	39.08 (4.72)	37.78 (6.16)
GSES[e]	32.41 (4.15)	31.37 (4.17)
Outcomes at one month		
PCS	50.74 (11.14)	48.20 (8.83)
MCS	47.57 (10.26)	45.83 (12.89)
HHI	39.45 (5.07)	38.61 (5.92)
GSES	32.99 (4.01)	30.02 (4.92)
Outcomes at three months		
PCS	50.25 (10.99)	50.37 (7.84)
MCS	47.48 (9.74)	44.69 (12.02)
HHI	39.89 (5.18)	38.52 (5.70)
GSES	32.76 (4.43)	29.21(6.19)

[a]MT4C: My Tools 4 Care.

[b]PCS: physical component score (SF-12v2).

[c]MCS: mental component score (SF-12v2).

[d]HHI: herth hope index.

[e]GSES: general self-efficacy scale.

Table 3. Analysis of covariance results for outcomes from baseline to three months for users versus nonusers (group).

Outcome	Parameter estimate (95% CI)	*P* value
SF-12v2[a] (PCS[b] and MCS[c]; n=76)		
PCS at three months		
Intercept	16.09 (8.60 to 23.58)	<.001
PCS—baseline	0.70 (0.56 to 0.83)	<.001
Group	−1.36 (−5.43 to 2.72)	.51
MCS at three months		
Intercept	16.77 (6.03 to 27.51)	.003
MCS at baseline	0.59 (0.39 to 0.79)	<.001
Group	2.79 (−2.35 to 7.93)	.28
HHI[d](n=78)		
HHI factor 1 at three months		
Intercept	2.16 (−0.48 to 4.80)	.11
HHI factor 1 at baseline	0.81 (0.61 to 1.00)	<.001
Group	0.59 (−0.35 to 1.53)	.22
HHI factor 2 at three months		
Intercept	4.21 (1.75 to 6.67)	.001
HHI factor 2 at baseline	0.67 (0.49 to 0.84)	<.001
Group	0.49 (−0.29 to 1.26)	.22
HHI factor 3 at three months		
Intercept	4.62 (2.09 to 7.15)	<.001
HHI factor 3 at baseline	0.62 (0.43 to 0.81)	<.001
Group	0.49 (−0.40 to 1.38)	.28
HHI total score at three months		
Intercept	5.50 (−1.62 to 12.62)	.13
HHI at baseline	0.84 (0.66 to 1.01)	<.001
Group	1.47 (−0.62 to 3.56)	.17
GSES[e] (n=77) at three months		
Intercept	4.14 (−2.59 to 10.86)	.22
GSES at baseline	0.78 (0.58 to 0.98)	<.001
Group	3.23 (1.12 to 5.33)	.003[f]

[a]SF-12v2: short-form 12-item health survey.

[b]PCS: physical component score (SF-12v2).

[c]MCS: mental component score (SF-12v2).

[d]HHI: herth hope index.

[e]GSES: general self-efficacy scale.

[f]Significant at $P<.05$.

Table 4. Generalized estimating equation results for outcomes (repeated measures analysis over three months [time 2=one month; time 3=three months]) for users compared with nonusers. Time 2 (one month from baseline) was not a significant factor in time 3 outcomes.

Outcome	Estimate	SE (95% CI)	*P* value
SF-12v2[a]			
PCS[b]			
Time 3[c]			
Group (users)	−2.03	1.68 (−5.32 to 1.27)	.23
Time 2	−0.23	0.94 (−2.07 to 1.60)	.80
MCS[d]			
Time 3[c]			
Group (users)	0.89	1.70 (−2.44 to 4.21)	.60
Time 2	−0.15	1.06 (−2.22 to 1.92)	.89
HHI[e] total			
Time 3[c]			
Group (users)	−0.38	0.86 (−2.06 to 1.30)	.66
Time 2	−0.44	0.52 (−1.45 to 0.58)	.40
GSES[f]			
Time 3[c]			
Group (users)	1.55	0.78 (0.01 to 3.09)	.048[g]
Time 2	0.06	0.42 (−0.77 to 0.89)	.89

[a]SF-12v2: short-form 12-item health survey.

[b]PCS: physical component score (SF-12v2).

[c]Reference group.

[d]MCS: mental component score (SF-12v2).

[e]HHI: herth hope index.

[f]GSES: general self-efficacy scale.

[g]Significant at $P<.05$.

Reasons for Nonuse of My Tools 4 Care

Reasons for nonuse of MT4C reported in the qualitative data included the following: caregiving demands; problems accessing and navigating the site; and preference for paper or in-person contact. Participants reported being consumed with the role of caregiving and as a result did not have enough time to use MT4C. As one participant said: "...and I got to admit that it was, uh, something that, uh, I didn't go onto too much, just strictly because of all the other things that were—were going on this past month." Those who were able to find a little bit of time to look at MT4C found they were quickly distracted by the care recipient and ultimately did not use it. One participant described having to stop using MT4C to tend to her husband: "Well, um, I just finished reading it, and—and—and, then, I had to go off because I had to go help my husband."

Caregiving demands also resulted in nonuser participants feeling stressed. As one participant said: "I'm extremely stressed with taking care of my wife, and so I lost the email with login instructions." Another indicated that the lack of energy was a factor "...[I] work full time early morning to late evening...and at the end of the day, I don't have the energy or time to go on the computer."

Problems accessing MT4C were related to poor internet connections, computer literacy, and difficulties navigating the site. Nonusers who lived in rural areas reported poor internet connections: "...my internet connection at home is poor—I live in a rural area." Several nonusers described their lack of experience with computers (computer literacy): "No, it—it's, uh, as far as the computer is concerned, it's the—the operator of it that's at fault."; another said: "Um, well, I get frustrated at myself when, you know, I'm working on the website..."

In terms of difficulty navigating the site, participants described forgetting the link to log in and difficulty printing instructions for the site. A participant described her frustration with not being able to find where she had previously been working on the site after being interrupted by her husband, "he kept interrupting me. Then, I couldn't find where I left off to continue..." Caregiving demands coupled with navigating the site were the reason this participant did not use MT4C.

Nonusers also seemed to have a preference for access to hard copy or paper format of MT4C and interaction with other carers. As a participant said, "...Sometimes, you actually have to have something printed in front of you, uh, and I'm better off—I'm better with paper. In some instances, to sit and reflect, I'm not really good at what—I'm not really one of those people who can do it all on-line." This participant described his lack of experience with working on the Web, but also suggested his preference was for paper. Another participant suggested a preference for social interaction rather than Web-based tools: "I think—I think I know—and this is [chuckles]—this isn't specific to this Toolkit, but it sort of relates to it: um, I think I'm the kind of person who gets a lot more out of, you know, actual social interaction around something."

The reasons for nonuse of MT4C, although reported as separate issues, appeared to be interrelated. For example, the lack of computer literacy meant it took more time to use MT4C, but with caregiving demands, less time was available. Furthermore, participants who are not computer literate became frustrated and thus preferred a paper format.

Discussion

General Self-Efficacy

The main findings of this study were a statistically significant difference between the users and nonusers of MT4C with regard to general self-efficacy (the confidence in their ability to deal with difficult situations). General self-efficacy significantly increased in the user group.

Other studies have found that general self-efficacy has a significant positive relationship with the quality of life of family carers of people with chronic illness [28]. For family carers of persons with dementia, self-efficacy has been found to have a significant relationship with health outcomes such as mental health [29,30] and, in particular, is negatively correlated with depression [31,32]. In the larger study, hope was found to be positively associated with general self-efficacy [13]. When the hope of participants increased, so did their confidence in their ability to deal with difficult situations. However, in the larger study using the intent-to-treat analysis, which included nonusers, general self-efficacy was not found to be significantly different in the treatment versus the control group, but hope was. It was important to further examine differences between users and nonusers in the treatment group, as the finding that self-efficacy was higher in users versus nonusers supports the intervention model in which MT4C has the potential to increase quality of life by increasing self-efficacy.

Reasons for Nonuse

The qualitative data suggested that caregiving demands, accessibility to the site, and preference for a paper version or face-to-face interaction were barriers to use for nonuser participants. Caregiving demands with subsequent family carer lack of energy and feelings of stress are consistent with the findings from the qualitative data from all participants in the larger study [19]. Quality of life scores at each time period for users and nonusers were not significantly different; however, whether the nonuser group experienced a more pronounced lack of energy and higher levels of stress than those in the user group is unclear. Future research should potentially also measure fatigue and stress as possible barriers to the use of Web-based interventions.

Poor connectivity to the internet was described as a barrier to use of MT4C by nonuser participants. Web-based interventions have been considered to be of benefit particularly for rural populations because of considerations related to accessibility [33]. However, poor connectivity to the internet, particularly for persons in rural areas in Canada, is a barrier to the use of any Web-based intervention [34,35]. Poor connectivity should be a concern for any research with Web-based interventions and possibly an exclusion criterion for participants in efficacy and effectiveness trials.

Computer literacy (ie, the ability to use computers and related technology efficiently) appeared to be one of the barriers to using MT4C. Inclusion criteria for the study included access to a computer and an email address. However, in this study, a measure of computer literacy was not used. Park et al [36], following an integrated review of health-related internet use of family carers of children, suggested that Web-based interventions should also include educational programs to increase computer literacy. Although MT4C was previously determined to be easy to use, an additional *tell me more* feature could be embedded into the program to assist carers who have low computer literacy.

What is unclear from our study is if access issues and computer literacy resulted in some participant preferences for a paper format and/ or face-to-face supportive interactions. Moreover, the nonusers referred in their comments to in-person interaction with other carers not Web-based interaction with other carers. When interaction with other family carers was added to a Web-based intervention, caregiving demand and computer literacy were also found to have an impact on the perceived benefit of a Web-based intervention [37]. This suggests that Web-based interventions to support family carers of persons with ADRD and MCC should not be the only format for support, but opportunities to use a paper version of MT4C and in-person face-to-face interactions are also important.

Limitations

This study was a secondary analysis; thus, follow-up interviews with nonuser participants were not conducted. Follow-up interviews would have been completed to further explore participants' reasons for nonuse and to answer questions about the relationship between computer literacy and their preferences for a paper format. Low computer literacy and poor connectivity were also not considered as exclusion criteria for the study, possibly influencing the results.

Importantly, this comparison involves a sample that was not randomly assigned, thus limiting the generalizability of the study. Although there were no statistically significant differences in the demographic characteristics between users and nonusers, there may be a potential imbalance between the groups based on unmeasured contextual characteristics. The findings contribute to the developing model of the intervention; however, in future studies, potential mediators and moderators should be

identified and their influence on the outcomes of the intervention should be evaluated.

Conclusions

The findings of this study reflect how comparisons of users and nonusers in Web-based intervention studies can improve Web-based interventions and the design of future studies. The statistically significant higher levels of general self-efficacy (or the confidence in the ability to deal with difficult situations) in users of MT4C is an important finding. Family carers of persons with dementia have reported significantly less self-efficacy than carers of persons without dementia [38]. As such, MT4C can potentially benefit family carers who are willing and able to use Web-based interventions. More research is needed to determine if adding an educational program for computer literacy may assist more family carers to access this Web-based intervention. In addition, future research should explore the use of MT4C in carers with diverse cultural backgrounds and languages.

Acknowledgments

This study is part of a program of research (Aging, Community and Health Research Unit) supported by the Canadian Institutes of Health Research Signature Initiative in Community-Based Primary Healthcare Funding (Reference Number: TTF 128261), MM-R's Canada Research Chairs Program, and WD's Endowed Research Chair University of Alberta. The authors wish to thank all of the study participants and their community partners: Alberta Caregivers Association, Alzheimer Society of Alberta and North West Territories, Alzheimer Society of Brant, Haldimand Norfolk, Hamilton and Halton, Alzheimer Society of Calgary, Alzheimer Society of Ontario, and Covenant Health Alberta. The authors also thank Tracy Chambers and the data collectors: Ann Cotton, Vanessa Hang, Melissa Northwood, Lori Pollard, and Tracy Wildeman.

Conflicts of Interest

None declared.

References

1. World Health Organization. 2012. Dementia: A Public Health Priority URL: https://apps.who.int/iris/bitstream/handle/10665/75263/9789241564458_eng.pdf;jsessionid=5634250F7FD71437CC5F71F333670A6D?sequence=1 [accessed 2019-03-13] [WebCite Cache ID 76qdBO8Ab]
2. Ploeg J, Northwood M, Duggleby W, McAiney CA, Chambers T, Peacock S, et al. Caregivers of older adults with dementia and multiple chronic conditions: exploring their experiences with significant changes. Dementia (London) 2019 Mar 6:1471301219834423 (epub ahead of print)(forthcoming)(forthcoming). [doi: 10.1177/1471301219834423] [Medline: 30841745]
3. Roth DL, Fredman L, Haley WE. Informal caregiving and its impact on health: a reappraisal from population-based studies. Gerontologist 2015 Apr;55(2):309-319 [FREE Full text] [doi: 10.1093/geront/gnu177] [Medline: 26035608]
4. Fonareva I, Oken BS. Physiological and functional consequences of caregiving for relatives with dementia. Int Psychogeriatr 2014 May;26(5):725-747 [FREE Full text] [doi: 10.1017/S1041610214000039] [Medline: 24507463]
5. Bangerter LR, Griffin J, Harden K, Rutten LJ. Health information–seeking behaviors of family caregivers: analysis of the health information national trends survey. JMIR Aging 2019 Jan 14;2(1):e11237. [doi: 10.2196/11237]
6. Guay C, Auger C, Demers L, Mortenson WB, Miller WC, Gélinas-Bronsard D, et al. Components and outcomes of internet-based interventions for caregivers of older adults: systematic review. J Med Internet Res 2017 Sep 19;19(9):e313 [FREE Full text] [doi: 10.2196/jmir.7896] [Medline: 28928109]
7. Sherifali D, Ali MU, Ploeg J, Markle-Reid M, Valaitis R, Bartholomew A, et al. Impact of internet-based interventions on caregiver mental health: systematic review and meta-analysis. J Med Internet Res 2018 Jul 3;20(7):e10668 [FREE Full text] [doi: 10.2196/10668] [Medline: 29970358]
8. Ploeg J, Ali MU, Markle-Reid M, Valaitis R, Bartholomew A, Fitzpatrick-Lewis D, et al. Caregiver-focused, web-based interventions: systematic review and meta-analysis (part 2). J Med Internet Res 2018 Oct 26;20(10):e11247 [FREE Full text] [doi: 10.2196/11247] [Medline: 30368439]
9. Chiu T, Marziali E, Colantonio A, Carswell A, Gruneir M, Tang M, et al. Internet-based caregiver support for Chinese Canadians taking care of a family member with Alzheimer disease and related dementia. Can J Aging 2009 Dec;28(4):323-336. [doi: 10.1017/S0714980809990158] [Medline: 19925698]
10. Lewis ML, Hobday JV, Hepburn KW. Internet-based program for dementia caregivers. Am J Alzheimers Dis Other Demen 2010 Dec;25(8):674-679. [doi: 10.1177/1533317510385812] [Medline: 21131674]
11. Ducharme F, Dubé V, Lévesque L, Saulnier D, Giroux F. An online stress management training program as a supportive nursing intervention for family caregivers of an elderly person. Can J Nurs Inform 2011;6(2):1-9 [FREE Full text]
12. Marziali E, Garcia LJ. Dementia caregivers' responses to 2 internet-based intervention programs. Am J Alzheimers Dis Other Demen 2011 Mar;26(1):36-43. [doi: 10.1177/1533317510387586] [Medline: 21282276]
13. Duggleby W, Swindle J, Peacock S. Self-administered intervention for caregivers of persons with Alzheimer's disease. Clin Nurs Res 2014 Mar;23(1):20-35. [doi: 10.1177/1054773812474299] [Medline: 23393291]
14. MyTools4Care. URL: https://www.mytools4care.ca/ [accessed 2019-07-01]

15. Duggleby W, Ploeg J, McAiney C, Peacock S, Fisher K, Ghosh S, et al. Web-based intervention for family carers of persons with dementia and multiple chronic conditions (My Tools 4 Care): pragmatic randomized controlled trial. J Med Internet Res 2018 Jun 29;20(6):e10484 [FREE Full text] [doi: 10.2196/10484] [Medline: 29959111]
16. Eysenbach G. The law of attrition. J Med Internet Res 2005 Mar 31;7(1):e11 [FREE Full text] [doi: 10.2196/jmir.7.1.e11] [Medline: 15829473]
17. Eysenbach G, Consort-eHealth Group. Consort-eHealth: improving and standardizing evaluation reports of Web-based and mobile health interventions. J Med Internet Res 2011 Dec 31;13(4):e126 [FREE Full text] [doi: 10.2196/jmir.1923] [Medline: 22209829]
18. Duggleby W, Ploeg J, McAiney C, Fisher K, Swindle J, Chambers T, et al. Study protocol: pragmatic randomized control trial of an internet-based intervention (My tools 4 care) for family carers. BMC Geriatr 2017 Aug 14;17(1):181 [FREE Full text] [doi: 10.1186/s12877-017-0581-6] [Medline: 28806917]
19. Ploeg J, McAiney C, Duggleby W, Chambers T, Lam A, Peacock S, et al. A web-based intervention to help caregivers of older adults with dementia and multiple chronic conditions: qualitative study. JMIR Aging 2018 Apr 23;1(1):e2. [doi: 10.2196/aging.8475]
20. Alzheimer Society of Canada. 2013. Progression of Alzheimer's Disease: Overview URL: https://alzheimer.ca/sites/default/files/files/national/progression-series/progression_overview_e.pdf [accessed 2019-03-13] [WebCite Cache ID 76qeUh3vE]
21. Herth K. Abbreviated instrument to measure hope: development and psychometric evaluation. J Adv Nurs 1992 Oct;17(10):1251-1259. [doi: 10.1111/j.1365-2648.1992.tb01843.x] [Medline: 1430629]
22. Schwarzer R. Self-Efficacy - Userpage. 2012. General Self-Efficacy Scale (GSE) URL: http://userpage.fu-berlin.de/health/selfscal.htm [accessed 2019-03-13] [WebCite Cache ID 76qTcWX3y]
23. Markowitz JS, Gutterman EM, Sadik K, Papadopoulos G. Health-related quality of life for caregivers of patients with Alzheimer disease. Alzheimer Dis Assoc Disord 2003;17(4):209-214. [doi: 10.1097/00002093-200310000-00003] [Medline: 14657784]
24. Arango-Lasprilla JC, Lehan T, Drew A, Moreno A, Deng X, Lemos M. Health-related quality of life in caregivers of individuals with dementia from Colombia. Am J Alzheimers Dis Other Demen 2010 Nov;25(7):556-561. [doi: 10.1177/1533317510382287] [Medline: 20826708]
25. Ware JE. User's Manual for the SF-12v2 Health Survey (with a Supplement Documenting SF-12 Health Survey). Lincoln, Rhode Island, United States: Quality Metric Incorporated; 2002.
26. Thabane L, Mbuagbaw L, Zhang S, Samaan Z, Marcucci M, Ye C, et al. A tutorial on sensitivity analyses in clinical trials: the what, why, when and how. BMC Med Res Methodol 2013 Jul 16;13:92 [FREE Full text] [doi: 10.1186/1471-2288-13-92] [Medline: 23855337]
27. Graneheim UH, Lundman B. Qualitative content analysis in nursing research: concepts, procedures and measures to achieve trustworthiness. Nurse Educ Today 2004 Mar;24(2):105-112. [doi: 10.1016/j.nedt.2003.10.001] [Medline: 14769454]
28. Duggleby W, Williams A, Ghosh S, Moquin H, Ploeg J, Markle-Reid M, et al. Factors influencing changes in health related quality of life of caregivers of persons with multiple chronic conditions. Health Qual Life Outcomes 2016 May 27;14:81 [FREE Full text] [doi: 10.1186/s12955-016-0486-7] [Medline: 27229926]
29. Zhang S, Edwards H, Yates P, Guo Q, Li C. Partial mediation role of self-efficacy between positive social interaction and mental health in family caregivers for dementia patients in Shanghai. PLoS One 2013;8(12):e83326 [FREE Full text] [doi: 10.1371/journal.pone.0083326] [Medline: 24386178]
30. Crellin NE, Orrell M, McDermott O, Charlesworth G. Self-efficacy and health-related quality of life in family carers of people with dementia: a systematic review. Aging Ment Health 2014;18(8):954-969 [FREE Full text] [doi: 10.1080/13607863.2014.915921] [Medline: 24943873]
31. Gallagher D, Mhaolain AN, Crosby L, Ryan D, Lacey L, Coen RF, et al. Self-efficacy for managing dementia may protect against burden and depression in Alzheimer's caregivers. Aging Ment Health 2011 Aug;15(6):663-670. [doi: 10.1080/13607863.2011.562179] [Medline: 21547745]
32. Rabinowitz YG, Saenz EC, Thompson LW, Gallagher-Thompson D. Understanding caregiver health behaviors: depressive symptoms mediate caregiver self-efficacy and health behavior patterns. Am J Alzheimers Dis Other Demen 2011 Jun;26(4):310-316. [doi: 10.1177/1533317511410557] [Medline: 21636582]
33. Deeken F, Rezo A, Hinz M, Discher R, Rapp MA. Evaluation of technology-based interventions for informal caregivers of patients with dementia - a meta-analysis of randomized controlled trials. Am J Geriatr Psychiatry 2019 Apr;27(4):426-445. [doi: 10.1016/j.jagp.2018.12.003] [Medline: 30642650]
34. Greenberg AJ, Haney D, Blake DK, Moser RP, Hesse BW. Differences in access to and use of electronic personal health information between rural and urban residents in the United States. J Rural Health 2018 Mar;34(Suppl 1):s30-s38 [FREE Full text] [doi: 10.1111/jrh.12228] [Medline: 28075508]
35. Federation of Canadian Municipalities. 2014. Broadband Access in Rural Canada: The Role of Connectivity in Building Vibrant Communities URL: https://fcm.ca/sites/default/files/documents/resources/report/broadband-access-rural-canada.pdf [accessed 2019-03-13] [WebCite Cache ID 76qTIUsqd]
36. Park E, Kim H, Steinhoff A. Health-related internet use by informal caregivers of children and adolescents: an integrative literature review. J Med Internet Res 2016 Mar 3;18(3):e57 [FREE Full text] [doi: 10.2196/jmir.4124] [Medline: 26940750]

37. Vaughan C, Trail TE, Mahmud A, Dellva S, Tanielian T, Friedman E. Informal caregivers' experiences and perceptions of a web-based peer support network: mixed methods study. J Med Internet Res 2018 Aug 28;20(8):e257 [FREE Full text] [doi: 10.2196/jmir.9895] [Medline: 30154074]
38. Son J, Erno A, Shea DG, Femia EE, Zarit SH, Stephens MA. The caregiver stress process and health outcomes. J Aging Health 2007 Dec;19(6):871-887. [doi: 10.1177/0898264307308568] [Medline: 18165286]

Abbreviations

ADRD: Alzheimer disease and related dementias
ANCOVA: analysis of covariance
GEE: generalized estimating equations
GSES: general self-efficacy scale
HHI: herth hope index
HRQOL: health-related quality of life
MCC: multiple chronic conditions
MCS: mental component score (SF-12v2)
MT4C: My Tools 4 Care
PCS: physical component score (SF-12v2)
SF-12v2: short-form 12-item health survey

Edited by G Eysenbach; submitted 03.04.19; peer-reviewed by F Lanfranchi, M Usman; comments to author 08.07.19; revised version received 08.08.19; accepted 18.08.19; published 17.10.19.

Please cite as:
Duggleby W, Ploeg J, McAiney C, Fisher K, Jovel Ruiz K, Ghosh S, Peacock S, Markle-Reid M, Williams A, Triscott J, Swindle J
A Comparison of Users and Nonusers of a Web-Based Intervention for Carers of Older Persons With Alzheimer Disease and Related Dementias: Mixed Methods Secondary Analysis
J Med Internet Res 2019;21(10):e14254
URL: https://www.jmir.org/2019/10/e14254
doi:10.2196/14254
PMID:31625947

Original Paper

Reduced Hospitalizations, Emergency Room Visits, and Costs Associated with a Web-Based Health Literacy, Aligned-Incentive Intervention: Mixed Methods Study

Jeffrey C Greene[1], BS; Jolie N Haun[2], EdS, PhD; Dustin D French[3], PhD; Susan L Chambers[4], MD; Robert H Roswell[5], MD

[1]MedEncentive, Oklahoma City, OK, United States

[2]College of Public Health, University of South Florida, Tampa, FL, United States

[3]Department of Ophthalmology, Feinberg School of Medicine, Northwestern University, Chicago, IL, United States

[4]Oklahoma City Gynecology & Obstetrics, Oklahoma City, OK, United States

[5]University of Oklahoma College of Medicine, Health Administration and Policy, University of Oklahoma College of Public Health, Oklahoma City, OK, United States

Corresponding Author:
Jeffrey C Greene, BS
MedEncentive
755 Research Parkway
Suite 440
Oklahoma City, OK,
United States
Phone: 1 405 627 9351
Email: jgreene@medencentive.com

Abstract

Background: The association between health literacy and health care costs, particularly for hospitalizations and emergency room services, has been previously observed. Health information interventions aimed at addressing the negative impacts of inadequate health literacy are needed. The MedEncentive Mutual Accountability and Information Therapy (MAIT) Program is a Web-based system designed to improve health and lower costs by aligning patient-doctor incentives.

Objective: In this mixed methods study of a Web-based patient-doctor aligned-incentive, information therapy program conducted in an 1800-member employee health plan, we aimed to (1) determine the program's quantitative impact on hospitalization and emergency room utilization and costs, and (2) assess survey responses about the program's perceived value.

Methods: We used a mixed methods, single within-group, pre-post, descriptive study design. We analyzed quantitative data using pre-post mean utilization and cost differences and summarized the data using descriptive statistics. We used open-ended electronic survey items to collect descriptive data and analyzed them using thematic content analysis.

Results: Hospitalizations and emergency room visits per 1000 decreased 32% (26.5/82.4) and 14% (31.3/219.9), respectively, after we implemented the program in 2015-2017, relative to 2013-2014. Correspondingly, the plan's annual per capita expenditures declined US $675 (95% CI US $470-865), or 10.8% ($675/$6260), after program implementation in 2015-2017 (US $5585 in 2013-2014 dollars), relative to the baseline years of 2013-2014 (US $6260; $P<.05$). Qualitative findings suggested that respondents valued the program, benefiting from its educational and motivational aspects to better self-manage their health.

Conclusions: Analyses suggested that the reported reductions in hospitalizations, emergency room visits, and costs were associated with the program. Qualitative findings indicated that targeted users perceived value in participating in the MAIT Program. Further research with controls is needed to confirm these outcomes and more completely understand the health improvement and cost-containment capabilities of this Web-based health information, patient-doctor, aligned-incentive program.

(J Med Internet Res 2019;21(10):e14772) doi:10.2196/14772

KEYWORDS

cost control; health care costs; health literacy; information therapy; aligned incentives; mutual accountability

Introduction

Background

Health literacy is defined as the "degree to which individuals have the capacity to obtain, process, and understand basic health information and services needed to make appropriate health decisions" [1] and "the capacity of individuals to obtain, interpret, and understand basic health information and services and the competence to use such information and services in ways which enhance health" [1]. The association between patients' health literacy levels and hospitalizations, preventable emergency room visits, and overall health care costs is established in the literature [2-7]. Not only is inadequate health literacy harmful and expensive, it is also prevalent. The 2003 National Assessment of Adult Literacy suggests that only 1 in 9 adults in the United States has proficient health literacy [6], contributing to billions of dollars in preventable expenditures per year [7]. Previous population-level studies have confirmed, when controlling for other person-level factors, that lower health literacy is a significant, independent factor associated with increased health care utilization and costs [2,3,5,8-10]. Specifically, inadequate health literacy has been associated with higher rates of hospitalizations and preventable emergency room visits [2,8,9]. Citing overwhelming empirical evidence, the US Department of Health & Human Services designated health literacy improvement as a top priority in 2010 [10]. Although inadequate health literacy is harmful, expensive, and prevalent, there have been few viable solutions to address the effects of inadequate health literacy in the general population and, much less, best practices to narrow the doctor-patient information asymmetry on a group level [5]. Best practices in the field of health literacy have recommended the need for universal precautions [11].

For decades, financial incentives to improve health care and health behaviors have been directed toward physicians and patients separately, with marginal success [12,13]. A recent study, in which a form of patient-doctor, aligned incentives was compared with traditional methods, found that the aligned-incentive approach produced superior outcomes [14]. Based on this finding, leading researchers in the field of behavioral economics concluded that "[aligned] financial incentives for patients and physicians could generate synergies that help patients, physicians, and health insurers achieve greater improvements in population health" [13]. Similar to health literacy, there are few, if any, viable patient-doctor, aligned-incentive solutions.

Information therapy is a term for "supplying patients with health information, enabling them to make informed decisions about their health and care, participate in their own well-being, and thus decrease the utilization of healthcare resources" [15]. It is further defined as providing patients with the right information, at the right time, in the right way, so patients can make informed decisions about their health [16]. Compensating physicians to provide an information therapy prescription to their patients as a reimbursable service is a concept suggested in the literature years ago but, heretofore, never attempted in a real-world setting [17]. Incentivizing patients to engage in information therapy and demonstrate assimilation of the information is a new concept, as is the idea of offering patient-doctor aligned incentives to empower and motivate patients with knowledge to self-manage their health. Reward-induced information therapy has the potential to offer a simple and sustainable solution to mitigating the debilitating effects of inadequate health literacy, in a manner that improves health and lowers per capita utilization and expenditures [17,18].

Prior research suggests that the majority of the general population have inadequate or marginal health literacy, requiring remediation when accessing the health care system [6]. Information therapy, as an established, systemwide practice, potentially provides a universal approach to support patients' health information needs—a key contribution to health literacy. We contend that utilization and cost are established outcome proxies to assess effects associated with information therapy [2,7].

Objective

The purpose of this study was to evaluate the outcomes of an employee health plan over a 5-year period, before and after the introduction of a Web-based information therapy, patient-doctor, aligned-incentive program. More specifically we aimed to (1) determine the impact (quantitatively) of the program on inpatient and emergency room utilization and costs, which also are proxies for overall health status, before and after implementation; and (2) evaluate participants' experiences using the program in correlation with the quantitative results.

The MedEncentive Mutual Accountability and Information Therapy Program

The MedEncentive Mutual Accountability and Information Therapy (MAIT) Program is a Web-based, mobile-enabled, information therapy, patient-doctor aligned-incentive program that promotes patient education and personal accountability, and supports health care cost containment [17,18]. MedEncentive's customers are health insurance plans sponsored by self-insured employers, governments, health systems, and commercial insurers. The MedEncentive program augments the sponsor's health plan (*plan*) as an additional benefit to the plan's members (*beneficiaries*). In the case of self-insured employers, the plan's summary plan description is modified to recognize the program as a benefit. As part of the service agreement, the plan sponsor directs its plan administrator (third-party administrator) to electronically transmit plan-member demographic enrollment and claims files to MedEncentive's computer system, and receive reward files for payment to doctors and patients who participate in the program. These electronic data exchanges employ industry-standard transmission protocols and data formats, so that they are secure, automated, and maintenance-free.

Program Overview

The MAIT Program uses plan sponsor-supplied member enrollment data to send orientation letters and personalized membership identification cards to all adult plan members. A program *opportunity*, for both doctors and patients, is initiated when doctors access the program's website, or as a result of MedEncentive's receipt of a claim associated with a covered

service. Covered services include any visit, consultation, or preventive examination rendered in-office to a covered member, by physicians (eg, doctor of medicine, doctor of osteopathy) of any medical specialty, or by licensed physician extenders (ie, nurse practitioners or physician assistants), for any medical condition or wellness examination.

The Provider's Experience

MedEncentive uses the diagnosis from the transmitted office visit claim to notify physicians of program opportunities via fax and email. These notices direct physicians to the MedEncentive website, where they may elect to participate in the program in 2 ways: (1) on a point-of-service (POS)-initiated basis, or (2) on a claims-initiated basis. The POS-initiated version is typically accomplished when practice personnel assign an identifier in their in-office computer system to patients covered by the program. These systems automatically notify doctors to access the program's website during or shortly after a covered office visit, to initiate an opportunity by entering the patient's diagnosis (see Multimedia Appendix 1). The MAIT Program can be integrated with in-office systems, which was the case in this implementation.

The claims-initiated version serves as a safety net in case a POS-initiated opportunity is missed. In the claims-initiated version, the MedEncentive system monitors incoming insurance claims to see whether physicians have previously used the POS-initiated version of the program. If not, then the system uses claim information to preload the patient's diagnosis and send the doctor the fax or email opportunity notice. When doctors choose to participate via the POS- or claims-initiated version, they access the program's website to complete 2 tasks: (1) consider evidence-based medicine treatment guidelines, and (2) select a patient educational article that the program's computer system lists in relevancy order to the patient's diagnosis (see Multimedia Appendix 2). The POS-initiated version functions identically to the claims-initiated version, with the exception of diagnosis input, time limits, and level of compensation. Since the program places a premium on timeliness, doctors earn US $15.00 for completing a POS-initiated session and US $7.50 for completing a claims-initiated session.

When physicians do not participate within 4 days of a claims-initiated notification, patients select their own articles from the list in conjunction with their program opportunity. As a result of this accommodation, both doctors and patients can earn the program's financial rewards independently of the other party's participation.

The Patient's Experience

Patient opportunities are initiated as a result of doctor participation, or as a result of office visit claims processing. Patients are notified of their opportunities to participate, by email notices or letters sent to their home, after each office visit (see Multimedia Appendix 3). Patients have 2 weeks to complete their information therapy sessions. For successfully completing a session, patients earn a financial reward, typically a refund of their office visit copay of US $15 or more. To earn their financial reward, patients access the program's website to (1) read the prescribed or self-selected educational article (see Multimedia Appendix 4), (2) demonstrate their understanding of the health information by passing an open-book test or declaring their comprehension (see Multimedia Appendix 5), (3) declare their adherence or provide a reason for nonadherence (see Multimedia Appendix 6), (4) agree to allow their physician to review their knowledge and adherence assessments (see Multimedia Appendix 7), and (5) rate how consistent their physician's care is to what they have just learned about recommended treatments (see Multimedia Appendix 8). Participation within required time frames is referred to as an information therapy *success*. When an opportunity expires without completion, this is referred to as a *miss*. Once a quarter, patients are given a second chance to complete the misses that occurred during the previous 90 days.

Patient Educational Content

While the program can be adapted to most Web-based educational content, MedEncentive used Healthwise articles in this implementation. Healthwise, Incorporated (Boise, ID, USA) is a conflict-free, nonprofit organization, nationally recognized for providing evidence-based, easy-to-understand health education at the fifth-grade reading level. They supply technology solutions that integrate with complex health information technology systems, with expert guidance on behavior change and shared decision making within the field of health care.

Methods

Design

This study used a mixed methods, single within-group, pre-post, descriptive study design to evaluate the MedEncentive MAIT Program. We used open-ended electronic survey items to collect descriptive data. Multimedia Appendix 9 shows the program study flowchart.

Setting

The study involved the employee health plan of a not-for-profit, acute-care general hospital (health plan sponsor) located in a semirural community in the south-central United States. The hospital is staffed by more than 1400 employees, with more than 100 physicians representing more than 30 specialties.

Sample

The study sample comprised the plan sponsor's employees and their covered dependents, to include spouses and children. The employees in the health plan were hospital and clinic personnel, including doctors, nurses, and other medical professionals, as well as administrative and support staff. No person in the covered population was excluded from this study or its analyses. We did not adjust to account for new hires or terminations.

Data Sources

The analysis of the program implementation relied on multiple sources of data and related background information. The health plan sponsor, its third-party administrator, and its pharmacy benefits manager were the primary sources of plan-member enrollment, medical claims, and pharmacy expenditure data,

from 2013 through 2017. We compiled the provider and patient program activity data and survey responses from MedEncentive's computer system.

Quantitative Procedures

Quantifying the impact of the MAIT Program involved a careful, step-by-step process of compiling and evaluating doctor and patient participation rates; the health plan's 2013-2017 hospitalizations, emergency room visits, and total expenditures; and other demographic and comorbid condition variables, before and after the program was implemented.

Doctor and Patient Participation Rates

The MedEncentive computer system automatically calculates doctor and patient participation using a unit of measure called *success rate*. This metric is derived by dividing the number of program successes (successfully completed information therapy sessions) by the total number of opportunities (office visits) incurred by all covered plan members. While patients need to know their doctors have an opportunity to participate—thus making physician inclusion in the program essential—patient success rate is the metric most aligned with reductions in hospitalizations and per capita expenditures at a group level.

Hospitalizations, Emergency Room Visits, and Total Health Care Expenditures

Detailed claims data for medical services, excluding pharmacy, were transmitted to MedEncentive on a monthly basis by the plan sponsor and its third-party administrator. Each claim contained more than 200 data elements, such as type of service, diagnosis, rendering provider, service location, gross charges, and net payments. The claims data included physician compensation and patient rewards associated with the program.

We sorted these data by date of service (end date) to organize the medical activity into the year services were rendered, from 2013 through 2017. We observed a typical 90-day run-out period for each year to capture the incurred charges in the year they occurred. We removed dental and optometry claims, since these services are not covered or directly affected by the program. To account for total expenditures, we added the plan members' direct out-of-pocket payments (copay amount, coinsurance, and deductible) to the amount paid by the plan sponsor. The clinical and economic analyses included all health plan enrollees, before and after implementation, regardless of program participation status.

Qualitative Procedures

We collected descriptive data about users' experiences with the MAIT Program through open- and closed-item electronic surveys. Administered at the conclusion of every information therapy session, for both doctors and patients, the surveys were voluntary and had no effect on the participants' financial rewards associated with the program. The open-ended items asked physicians and patients about their experience with the program and for their suggestions for improvement. Closed-ended survey items asked patients about (1) how helpful the educational article was for managing their diagnosis or in maintaining their health (2) how closely they were following the health recommendations contained in this article, (3) how much physician access to the program's survey responses motivated them to improve their health literacy and health behaviors, (4) the importance of their physician's awareness regarding the patient's capacity for self-management, and (5) the importance of their physician's awareness of the patient's intention to accomplish health objectives.

Analysis

Quantitative Analysis

Since voluntary participation in the program by doctors and patients can be associated with the intended clinical and economic outcomes, we began our quantitative analysis by examining the standard doctor and patient success rate reports generated by the MedEncentive computer system. We were particularly interested in determining whether we had achieved the 55% patient success rate threshold, since it is predictive of the clinical and economic group-level effectiveness of the program.

We analyzed clinical and economic outcomes by comparing annual hospitalizations per 1000 enrollees, emergency room visits per 1000 enrollees, total expenditures per capita, and other variables for 2015-2017 (the implementation period) versus the baseline years (2013-2014), prior to implementing the program. We conducted pre-post analysis of mean cost differences, with confidence intervals, for emergency room, hospitalization, and total care costs [19-22].

Since the program was designed to motivate adherence to recommended treatments and mitigate the effects of inadequate health literacy, which are associated in the literature with hospitalizations, emergency room visits, and total expenditures, these measures were the most effective means to measure program effectiveness [23-26]. To compare annual preimplementation versus postimplementation per capita expenditures, we multiplied the post period (2015-2017) annual episodes (units) of care (hospitalizations, emergency room visits, outpatient services, and pharmacy scripts) by the annual unit costs incurred during the baseline period (2013-2014). The normalization adjustments were made in consultation with the health plan sponsor (hospital) to adjust for known variables, such as pricing, coding, and charge capture. There were no significant benefit design changes over the 5-year period. We considered other health improvement and cost-containment initiatives and, upon analysis, ruled them out as significant contributors to the outcomes analyzed (hospitalizations and emergency room utilization, and total costs).

Qualitative Analysis

Structured, open-ended survey-item data were deidentified and cleaned, and prepared for analysis. We managed qualitative survey data responses using thematic content analysis, based on the topic addressed in the structured items and the response (eg, adherence, program satisfaction). We analyzed data in 2 stages to identify domains and taxonomies related to participant experiences [27]. The first round of coding included summarizing and reducing data into preliminary metadomains. Methods included deductive structural coding and inductive descriptive coding, based on themes that emerged from the surveys. In a second round of coding, we reduced coded data

into meaningful domains. As we developed coding schemas to create domains, data samples were extracted and coded by at least two team members and evaluated for interrater reliability and validity.

Ultimately, we sorted open-ended survey data into program, provider (doctor), health status, and other categories. Comments pertaining to the program were coded as testimonials, suggestions and service requests, and complaints. Comments pertaining to providers were coded as testimonials and complaints. Comments related to the patient's health status were coded as general medical condition, improving medical condition, and worsening medical condition. Other comments included insurance complaints.

Results

Quantitative Sample-Based Findings

Table 1 presents the health plan's total enrollees per annum, the mean annual enrollment, and the number of enrollees who received health care, as well as demographic variables, for the years 2013 through 2017. This study's sample of patients comprised the plan sponsor's employees and their covered dependents (ie, spouses and children), with a mean of 1803 per year over the 5-year study period (Table 1). It is notable that the plan grew from a mean of 1660 enrollees in 2013 to 1960 enrollees in 2017. The number of plan members receiving care grew from 1560 in 2013 to 1863 in 2017, and the total number of plan members enrolled at any point during a calendar year grew from 1752 in 2013 to 2554 in 2017. This growth was due, in large part, to the expansion of the hospital's services and acquisition of local medical clinics.

Patient Success (Participation) Rates

Patient participation exceeded the targeted 55% success rate threshold in the first year, reaching 68.67% (4245/6182) at the end of 2015. Patient success rate continued to climb to 74.23% (5108/6881) in the first quarter of 2018 (see Multimedia Appendix 10). This level of patient participation predicted, with a high probability, that the clinical and economic outcome objectives would be achieved.

As Table 2 shows, young adults (18-29 years) had the lowest patient success rate (1885/2969, 63.49%) over the 3-year intervention period, while senior adults (≥65 years) had the highest success rate (697/894, 78.0%). Also, the 60- to 65-year age group had the greatest improvement in success rate, climbing 18.4% from 2015 to 2017. This suggests that those with the greatest need participated in the program most frequently and dispels the notion that older adults may be technology challenged or averse to Web-based health literacy mechanisms. Also notable, males and females participated at essentially the same rate (4544/6463, 70.31% vs 8700/12,212, 71.24%, respectively), though females made far more office visits, and the number of office visits per capita remained consistent over the 3-year period (3.3-3.4 visits per annum).

Provider Success (Participation) Rates

The overall annual provider success rate started at 30.62% (1890/6173) at the end of 2015 and climbed to 45.41% (2619/5768) by the end of 2016 (see Multimedia Appendix 10). The providers employed by the hospital (health plan sponsor) achieved an even higher level of provider engagement, reaching 55.34% (1654/2989) by the end of 2016. This was due to two developments during the first 18 months of implementation. First, the hospital (plan sponsor) changed its policy of retaining the program's compensation earned by its employed providers, agreeing to pass these payments on to their participating physicians. Second, the hospital integrated the program with the hospital's electronic health record system, NextGen, a leading electronic health record supplier. This project was completed in 2016. As a result, doctors were able to access the program through the hospital's electronic health record system, while having patient demographic and diagnosis information directly transmitted to the MedEncentive computer system. Providers were also automatically notified of covered patients before, during, and immediately after an office visit, offering physicians greater opportunity to use the more timely and higher-paying real-time version of the program.

Clinical and Economic Outcomes

Program effectiveness can, in part, be measured by an improvement in clinical outcomes, including overall hospitalization and emergency room visit rates. We compiled the total hospitalizations and emergency room visits from the claims data for the 2 years prior to program implementation (2013-2014), and the 3 years after deployment (2015-2017), as Table 3 shows. All members enrolled in the health plan were included in the totals. As Multimedia Appendix 11 illustrates, 2013 and 2014 hospitalizations per 1000, before the introduction of the program, were 87.3 and 82.4, respectively. In 2015, 2016, and 2017, after program implementation, the hospitalization rates were 57.2, 53.9, and 56.6, respectively (P<.05). On average, this represents a 32% (26.5/82.4) decrease in admissions per 1000, relative to the baseline year of 2014. Emergency room visits per 1000 in 2013 and 2014 were 251.8 and 219.9, respectively, whereas in 2015, 2016, and 2017, emergency room visits per 1000 plan members decreased to 191.3, 187.5, and 187.3, respectively (P<.05). In summary, hospitalizations and emergency room visit rates per 1000 decreased 32% (26.5/82.4) and 14% (31.3/219.9), respectively, in 2015-2017 after implementation of the program, relative to 2013-2014, prior to program implementation, inclusive of all enrollees (participants and nonparticipants).

Table 1. Demographic characteristics of patients from 2013-2014 (before program implementation) and 2015-2017 (after program implementation).

Variables	Preimplementation		Postimplementation		
	2013	2014	2015	2016	2017
Enrollment, n					
Total receiving care[a]	1560	1619	1609	1729	1863
Total enrollees during year[b]	1752	1819	2205	2265	2554
Mean annual enrollment[c]	1660	1760	1783	1856	1960
Link to employee-based enrollment, n (%)					
Self[b]	744 (42.5)[c]	780 (42.9)	943 (42.9)	964 (42.6)	1065 (41.70)
Dependent[b]	1008 (57.53)[c]	1039 (57.12)	1262 (57.23)	1301 (57.44)	1489 (58.30)
Sex, n (%)					
Male	748 (42.7)[b,c]	776 (42.7)[a]	944 (42.9)[a]	965 (42.6)[a]	1103 (43.19)[a]
Female	1004 (57.30)[b,c]	1043 (57.30)[a]	1261 (57.20)[a]	1300 (57.40)[a]	1451 (56.81)[a]
Age group (years)[b], n (%)					
0-17	454 (25.9)[c]	468 (25.7)	542 (24.6)	560 (24.7)	630 (24.7)
18-29	304 (17.4)[c]	350 (19.2)	474 (21.5)	486 (21.5)	582 (22.8)
30-39	295 (16.9)[c]	298 (16.4)	377 (17.1)	373 (16.5)	392 (15.3)
40-49	257 (14.7)[c]	277 (15.1)	308 (14.0)	339 (15.0)	432(16.9)
50-59	274 (15.7)[c]	276 (15.2)	313 (14.2)	313 (13.8)	302 (11.8)
60-64	104 (6.0)[c]	102 (5.6)	121 (5.5)	123 (5.4)	142 (5.6)
65	62 (4)[c]	50 (3)	70 (3)	71 (3)	74 (3)
Age (years), mean[b]	33.0[c]	32.8	32.6	32.7	32.5
Office visits per person per year, mean[b]	2.9	3.0	2.9	3.0	3.2

[a]Total members treated during the year.

[b]Total plan-member enrollees during the year.

[c]2013 total enrollees extrapolated from members treated.

Table 2. Total patient success rate, total office visits, and overall success percentage (2015-2017), and 3-year trend by demographic characteristics.

Demographic variable	Patient success (n)	Total office visits (n)	Overall success (%)	Change 2015-2017 (%)
Age group (years)				
0-17	2876	3980	72.26	2.23
18-29	1885	2969	63.49	9.81
30-39	2235	2987	74.82	1.60
40-49	2125	2987	71.15	7.82
50-59	2218	3054	72.63	1.12
60-64	1208	1804	66.96	18.39
≥65	697	894	78.0	–6.0
Age total	13,244	18,675	70.92	4.88
Relationship				
Employee	6771	9361	72.33	4.05
Dependent	6473	9314	69.50	5.66
Relationship total	13,244	18,675	70.92	4.88
Sex				
Male	4544	6463	70.31	1.93
Female	8700	12,212	71.24	6.46
Sex total	13,244	18,675	70.92	4.88

Table 3. Hospitalizations and emergency room visits from 2013-2014 (before program implementation) and 2015-2017 (after program implementation).

Variables	Preimplementation		Postimplementation			
	2013	2014	2015	2016	2017	Postperiod mean
Mean annual enrollment, n	1660	1760	1783	1856	1960	1866
Hospital admissions, n	145	145	102	100	111	104
Admissions per 1000, n	87.3	82.4	57.2	53.9	56.6	55.9
Admissions per 1000 change from 2014, %	N/A[a]	N/A	–30.6	–34.6	–31.2	–32.1
Emergency room visits, n	418	387	341	348	367	352
Emergency room visits per 1000, n	251.8	219.9	191.3	187.5	187.3	188.6
Emergency room visits per 1000 change from 2014, %	N/A	N/A	–13.0	–14.7	–14.8	–14.2

[a]N/A: not applicable.

As Multimedia Appendix 12 illustrates, the plan's annual per capita expenditures, inclusive of all program costs, declined US $675 (95% CI US $470-865), or 10.8% ($675/$6260), after program implementation in 2015-2017 (US $5585 in 2013-2014 dollars), relative to the baseline years of 2013-2014 (US $6260; $P<.05$), inclusive of all enrollees (participants and nonparticipants) (see Table 4).

Table 4. Health care costs from 2013-2014 (before program implementation) and 2015-2017 (after program implementation).

Variables	Preimplementation		Postimplementation			
	2013	2014	2015	2016	2017	Postperiod mean
Total expenditures (all medical and pharmacy)[a], US $	9,940,434	11,468,059	9,462,011	10,580,146	10,726,060	10,256,072
Total program costs[b], US $	N/A[c]	N/A	156,403	165,577	179,207	167,062
Annual mean enrollment, n	1660	1760	1783	1856	1960	1866
Costs per member per year, US $						
Total expenditures without program costs[a]	5988	6516	5307	5701	5472	5495
Total program costs[b]	N/A	N/A	88	89	91	90
Total expenditures with program costs[a]	5988	6516	5395	5790	5564	5585
Mean baseline (2013-2014) expenditures	N/A	N/A	6260	6260	6260	6260
Gross savings[a,d]	N/A	N/A	953	559	787	764
Net savings[a,e]	N/A	N/A	865	470	696	675
Net savings, %	N/A	N/A	13.8	7.5	11.1	10.8

[a]2015-2017 amounts adjusted to 2013-2014 basis.

[b]N/A: not applicable.

[c]Total program costs for 2015-2017 include all patient rewards, physician compensation, and program administration fees.

[d]Gross savings for 2015-2017=2015-2017 average expenditures less program costs – 2013-2014 average expenditures.

[e]Net savings for 2015-2017=2015-2017 average expenditures with program costs – 2013-2014 average expenditures.

Quantitative Survey-Based Findings

During 2015-2017, the participating health plan members rated the helpfulness of the program's educational content at 4.40 out of 5 (with 5 being most helpful), representing 15,260 responses. These ratings indicated there was a strong consensus among patients that the program's educational content was helpful in managing their disease or condition, or in maintaining good health.

When patients were asked to report (to their doctors) their level of adherence with the health recommendations contained in the program's educational content, on a scale of 1 to 5, with 1 meaning "not following recommendations" and 5 meaning "following recommendations closely," the mean response was 4.70 (n=15,186) over the 2015-2017 time period. These self-assessments indicated a strong consensus among patients that they were, or intended to be, compliant with recommended treatments. Table 5 presents survey-item results.

Table 5. Patient responses to 5-point Likert-type scale survey items.

Survey item	Response option					Mean
	1	2	3	4	5	
How helpful has this article's information been to you in managing your disease or condition, or in maintaining your good health? n (%)	187 (1.2)	256 (1.7)	1511 (9.95)	4575 (30.13)	8731 (57.49)	4.40
Please share with your doctor how closely you are following the health recommendations contained in this article as you understand them. n (%)	42 (0)	29 (0)	155 (1.0)	3972 (26.03)	10,988 (72.01)	4.70

The survey item reflecting physician influence on the patient's motivation to gain health knowledge and improve health behaviors had a mean score of 8.80 out of 10, representing 13,401 responses, indicating a strong consensus that physicians positively influenced patients to improve their health literacy and health behaviors. When asked whether it was important for his or her physician to know of the patient's competency to self-manage, patient responses had a mean score of 9.24 out of 10, representing 13,401 responses. This indicated that the majority of patients thought it important that their doctor was aware that they understood how to manage their health. Finally, when asked whether it was important that their physicians knew that they were accomplishing health objectives, patient responses had a mean score of 9.26 out of 10, representing 13,401 responses. This indicated that the majority of patients thought it important that their doctor was aware that they were trying to accomplish health objectives. Table 6 presents survey response distributions for these items.

Table 6. Patient participant responses to 10-point Likert-type scale survey items.

Survey item	Response level (10=Most)										Mean
	10	9	8	7	6	5	4	3	2	1	
…how much does the knowledge that your physician has access to your questionnaire responses motivate you to improve your health literacy and health behaviors? n (%)	7419 (55.4)	2163 (16.1)	1580 (11.8)	653 (4.9)	552 (4.1)	567 (4.5)	71 (0.5)	99 (0.7)	61 (0.5)	197 (1.8)	8.8
…how important is it to you that your doctor is aware that you understand how to self-manage your health? n (%)	8488 (63.3)	2355 (17.6)	1376 (10.3)	495 (3.7)	290 (2.2)	250 (1.9)	34 (0.3)	30 (0.2)	21 (0.2)	62 (0.5)	9.2
…how important is it to you that your doctor is aware that you are trying to accomplish or are accomplishing health objectives? n (%)	8599 (64.2)	2334 (17.4)	1337 (10.0)	455 (3.4)	280 (2.1)	245 (1.8)	26 (0.2)	39 (0.3)	24 (0.2)	62 (0.5)	9.3

Qualitative Survey-Based Findings

During 2016-2017, patients posted 555 comments, with 323 (58.2%) pertaining to the program, 183 (33.0%) pertaining to their provider, 31 (5.6%) pertaining to health-related topics, and 18 (3.2%) pertaining to other general topics. Of the program-related comments, the majority (210/323, 65.0%) were testimonials, 33.4% (n=108) were suggestions or requests to improve the program, and 1.6% (n=5) were complaints. Table 7 lists exemplar patient and provider comments about the program.

Table 7. Exemplar patient and provider information therapy program-related comments.

Respondent type	Exemplar comments
Patient	Program is great – [puts] a good emphasize on personal accountability.
	Thank you for this MedEncentive program for us. It helps me a lot to know the cause, symptoms, prevention, medicine etc. of my illness. I appreciate it, that my employer has this kind of medical program for their employees.
	I like well enough that I come back later and read related articles.
	I think the more information the patient has the better. You're taking the right approach in educating the patient. Thank you!
	Excellent plan to get people to take ownership of their health status.
	This is [a] great incentive for patients to not only be aware of their own health and medication issues, but also an opportunity for them to read and learn more about these issues through a formal method of information retrieval. Many adults (old and young) rely upon internet to diagnose and learn about health issues and medications. This format is associated with health professionals and would seem to contain more reliable information and there is financial incentive to completion.
	I love the program. It is educational and beneficial. The financial incentive helps our family greatly as we use it for copays and supplies. We are very thankful for the program.
	I read articles on weight management, and boosting metabolism through exercise and dietary control. I found the sections on boosting metabolism by exercising vigorously at least 2.5 hours weekly in suggested increments, and the reasons for that to be interesting. I feel this is valuable, because I intend to bring my weight under better control.
	Good reminder. It's hard to remember everything discussed at the appt. this gives me a refresher to read on my own time.
Provider	Keeps me on my toes to talk to patients about diet and exercise.
	The educational handouts are easy to read and are very helpful for my patients. It helps me educate my patients.
	...good selection of articles, easy to prescribe, keep up the good work.
	This is one of the best and most expedient ways of reinforcing discussions we have with our patients in the office setting.

Discussion

Principal Findings

The MedEncentive MAIT Program is, to our knowledge, one of the first population-level solutions to use a Web-based approach, combining doctor and patient aligned incentives and information therapy, aimed at improving health and lower costs. Our analysis suggests that the MedEncentive MAIT Program was associated with meaningful reductions in health care utilization that were sustained into the third year of program implementation, through reductions in per capita expenditures and hospitalizations and emergency room use. These findings are in line with previous research on aligned-incentive program expenditure outcomes [18].

Our quantitative analysis found that, from 2015 through 2017, after the program was introduced, hospitalizations and emergency room visits per 1000 plan members, and per capita expenditures, declined relative to the 2013-2014 preimplementation period, by 32.1%, 14.2%, and 10.8%, respectively. Correspondingly, the qualitative survey items

suggested that the majority of respondents found the program's educational content to be very helpful in managing their disease or condition, and maintaining good health. Since patient adherence is such an important predictor of health status, service utilization, and costs [28], it is compelling that the surveys indicated that most patients intended to be compliant with recommended treatments. Furthermore, the surveys found that physicians positively influenced patients to improve their health literacy and health behaviors.

These findings indicated that it was important to patients for their doctor to be aware that they (1) understand how to self-manage their health, and (2) are trying to accomplish health objectives. These reported attitudes help explain the quantitative outcomes and set the stage for aligned-incentive programs to leverage information therapy as a means to mitigate the impact of inadequate health literacy. It is notable that other health improvement initiatives were launched by the plan sponsor, principally in the 2016-2017 time frame. However, the most significant improvement in hospitalizations and emergency room rates was temporally associated with the introduction of the program in 2015.

Programmatic Implications

Program adoption and retention are predicated on its ease of implementation and maintenance. Over the test period, the plan sponsor experienced continuous program access 99.8% of the time, with no reports of degradation in the program's website performance due to scaling or spikes in activity. The plan sponsor integrated the program with its clinic electronic health record system to streamline the provider experience. This integration functioned without difficulty throughout program implementation.

The effectiveness of any wellness, prevention, or managed-care program relies on high levels of patient and medical provider engagement in aspects of the program designed to improve health and health care. As the program participation statistics indicate, this goal was achieved and sustained, aided by the collaborative efforts of the plan sponsor and MedEncentive.

Study Limitations

Study limitations should be considered when interpreting our findings. First, though these findings are compelling, they can only be generalized to not-for-profit, acute-care, general-hospital employee health plans, located in the south-central United States. Second, the primary limitation is the internal nature of the evaluation. Future research should focus on external review and validation. Third, while this study offers associated confirmation of the program's effectiveness, it is not a randomized control trial and, therefore, falls short of the gold standard for determining causation; hence, further research is needed. Fourth, the conservative analysis may be considered a limitation, compared with more complex analyses (eg, quasi-experimental designs); however, this was prohibited so as to remain compliant with institutional review board standards. Fifth, this evaluation did not control for individual health literacy, but we contend that, statistically, the majority of the general population have inadequate or marginal health literacy and require health information support, such as information therapy; therefore, this program supports health literacy needs at the universal level and measures relevant and meaningful associated outcomes. Future research should control for health literacy. Sixth, we made adjustments to normalize postimplementation expenditures to baseline levels in consultation with the plan sponsor. These adjustments reflect known variances in medical services coding, pricing, charge capture, and benefit design. The precision of these adjustments is difficult to judge, which is further justification to test the program's capabilities by means of randomized control trials. Finally, although this was an opt-in study design, loss to follow-up can be an issue in cohort studies. Our study, however, included patients who used care throughout the study time frame.

Conclusions

Use of the Web-based MedEncentive MAIT Program was associated with a reduction in hospitalizations (26.5/82.4, 32%) and emergency room visits (31.3/219.9, 14%) per 1000 members. The plan's annual per capita expenditures declined US $675 (95% CI US $470-865), or 10.8% ($675/$6260), after program implementation in 2015-2017 (US $5585 in 2013-2014 dollars), relative to the baseline year of 2014 (US $6260; $P<.05$). Our qualitative analysis of participant survey responses corroborated these findings. Therefore, it is reasonable to conclude that the effectiveness of the program was evident in this study. Findings warrant investment in larger, longer-running randomized control trials to further examine and validate these results.

Acknowledgments

The authors would like to acknowledge Stillwater Medical Center, whose leadership and staff made this project possible. In particular, we extend our sincere appreciation to Keith Hufnagel, Vice President, Human Resources; Leva Swim, PhD, Vice President, Population Health; and Wendi Barrett, CBO Director, for their generosity of time and consultation throughout this project.

Conflicts of Interest

The authors respectfully wish to report potential conflicts of interest represented by financial benefit. JCG is the cofounder, chief executive officer, and part owner of MedEncentive, LLC, and receives no compensation from MedEncentive except for supplemental health insurance, valued at less than US $1000 per month. JCG is the inventor of the MedEncentive Mutual Accountability and Information Therapy Program's patented process, trademarked as the Trilateral Health Accountability Model, in which he holds a royalty interest. SLC is cofounder and part owner of MedEncentive, LLC. She is also a royalty holder in the Company's inventions. Both JCG and SLC are board members of MedEncentive, but neither is compensated in this role. As part owners and royalty holders, and in their aforementioned roles with the Company, JCG and SLC could financially benefit from the publication of this paper.

JNH and DDF received a consultation fee to provide expertise and support in the analysis, interpretation of data findings, and development of this manuscript. JNH and DDF do not have equity in the Company, nor is their payment contingent on the success of the Company; they do not sit on board or committee for the Company.

Multimedia Appendix 1
Patient's diagnosis entered in the program's website by the doctor.
[PDF File (Adobe PDF File), 86 KB - jmir_v21i10e14772_app1.pdf]

Multimedia Appendix 2
Doctors select relevant education for their patients.
[PDF File (Adobe PDF File), 104 KB - jmir_v21i10e14772_app2.pdf]

Multimedia Appendix 3
Patients are notified of their "opportunity" to earn a financial reward for participating in the program.
[PDF File (Adobe PDF File), 585 KB - jmir_v21i10e14772_app3.pdf]

Multimedia Appendix 4
Patients read educational article specific to their health.
[PDF File (Adobe PDF File), 157 KB - jmir_v21i10e14772_app4.pdf]

Multimedia Appendix 5
Patients answer questions to confirm their understanding of how to self-manage their health.
[PDF File (Adobe PDF File), 98 KB - jmir_v21i10e14772_app5.pdf]

Multimedia Appendix 6
Patients declare their adherence with recommended treatments, or provide a reason for nonadherence.
[PDF File (Adobe PDF File), 107 KB - jmir_v21i10e14772_app6.pdf]

Multimedia Appendix 7
Patients allow their physicians access to their knowledge assessment and adherence declaration.
[PDF File (Adobe PDF File), 83 KB - jmir_v21i10e14772_app7.pdf]

Multimedia Appendix 8
Patients rate how consistent their physician's care is to recommended treatments.
[PDF File (Adobe PDF File), 100 KB - jmir_v21i10e14772_app8.pdf]

Multimedia Appendix 9
MAIT Program study flow diagram.
[PDF File (Adobe PDF File), 61 KB - jmir_v21i10e14772_app9.pdf]

Multimedia Appendix 10
Patient and provider success rates.
[PDF File (Adobe PDF File), 128 KB - jmir_v21i10e14772_app10.pdf]

Multimedia Appendix 11
Hospitalization and emergency room rates.
[PDF File (Adobe PDF File), 158 KB - jmir_v21i10e14772_app11.pdf]

Multimedia Appendix 12
Economic outcomes of the health plan, pre- and post-MAIT Program implementation.
[PDF File (Adobe PDF File), 165 KB - jmir_v21i10e14772_app12.pdf]

References

1. Ratzan SC, Parker RM. Introduction. In: Selden CR, Zorn M, Ratzan SC, Parker R, editors. Current Bibliographies in Medicine: Health Literacy. Bethesda, MD: National Institutes of Health, U.S. Department of Health and Human Services; 2000:v.
2. Haun JN, Patel NR, French DD, Campbell RR, Bradham DD, Lapcevic WA. Association between health literacy and medical care costs in an integrated healthcare system: a regional population based study. BMC Health Serv Res 2015;15:249 [FREE Full text] [doi: 10.1186/s12913-015-0887-z] [Medline: 26113118]
3. Howard DH, Gazmararian J, Parker RM. The impact of low health literacy on the medical costs of Medicare managed care enrollees. Am J Med 2005 Apr;118(4):371-377. [doi: 10.1016/j.amjmed.2005.01.010] [Medline: 15808134]
4. Andersen R, Newman JF. Societal and individual determinants of medical care utilization in the United States. Milbank Q 2005 Nov 18;83(4). [doi: 10.1111/j.1468-0009.2005.00428.x]
5. Nielsen-Bohlman L, Panzer A, Kindig D. Health Literacy: A Prescription to End Confusion. Washington, DC: The National Academies Press; 2004.
6. Kutner M, Greenberg E, Jin Y, Paulsen C. The Health Literacy of America's Adults: Results From the 2003 National Assessment of Adult Literacy (NCES 2006-483). Washington, DC: National Center for Education Statistics; 2006 Sep. URL: https://nces.ed.gov/pubs2006/2006483.pdf [accessed 2019-09-05]
7. Health Policy Institute. Low Health Literacy Skills Increase Annual Health Care Expenditures by $73 Billion. Washington, DC: Georgetown University, McCourt School of Public Policy, Health Policy Institute; 1999 May 14. URL: https://hpi.georgetown.edu/healthlit/# [accessed 2918-09-05]
8. Howard DH, Sentell T, Gazmararian JA. Impact of health literacy on socioeconomic and racial differences in health in an elderly population. J Gen Intern Med 2006 Aug;21(8):857-861. [doi: 10.1111/j.1525-1497.2006.00530.x] [Medline: 16881947]
9. Baker DW, Parker RM, Williams MV, Clark WS, Nurss J. The relationship of patient reading ability to self-reported health and use of health services. Am J Public Health 1997 Jun;87(6):1027-1030. [Medline: 9224190]
10. U.S. Department of Health and Human Services. Healthy People 2010: Understanding and Improving Health. 2nd edition. Washington, DC: US Government Printing Office; 2000 Nov. URL: https://www.healthypeople.gov/2010/document/pdf/uih/2010uih.pdf [accessed 2019-09-11]
11. Brown DR, Ludwig R, Buck GA, Durham D, Shumard T, Graham SS. Health literacy: universal precautions needed. J Allied Health 2004;33(2):150-155. [Medline: 15239414]
12. Lorincz IS, Lawson BCT, Long JA. Provider and patient directed financial incentives to improve care and outcomes for patients with diabetes. Curr Diab Rep 2013 Apr;13(2):188-195 [FREE Full text] [doi: 10.1007/s11892-012-0353-9] [Medline: 23225214]
13. Navathe AS, Emanuel EJ, Volpp KG. Aligning patient and physician incentives. JAMA 2018 Oct 23;320(16):1635-1636. [doi: 10.1001/jama.2018.11245] [Medline: 30422278]
14. Asch DA, Troxel AB, Stewart WF, Sequist TD, Jones JB, Hirsch AG, et al. Effect of financial incentives to physicians, patients, or both on lipid levels: a randomized clinical trial. JAMA 2015 Nov 10;314(18):1926-1935 [FREE Full text] [doi: 10.1001/jama.2015.14850] [Medline: 26547464]
15. Mitchell DJ. Toward a definition of Information Therapy. Proc Annu Symp Comput Appl Med Care 1994:71-75 [FREE Full text] [Medline: 7950018]
16. Kemper DW, Mettler M. Information Therapy: Prescribed Information as a Reimbursable Medical Service. 1st edition. Boise, ID: Healthwise, Inc; 2002.
17. Keene N, Chesser A, Hart TA, Twumasi-Ankrah P, Bradham DD. Preliminary benefits of information therapy. J Prim Care Community Health 2011 Jan 01;2(1):45-48. [doi: 10.1177/2150131910385005] [Medline: 23804662]
18. Parke DW. Impact of a pay-for-performance intervention: financial analysis of a pilot program implementation and implications for ophthalmology (an American Ophthalmological Society thesis). Trans Am Ophthalmol Soc 2007;105:448-460 [FREE Full text] [Medline: 18427625]
19. Rothman KJ, Greenland S, Lash TL. Modern Epidemiology. Volume 3. Philadelphia, PA: Wolters Kluwer Health/Lippincott Williams & Wilkins; 2008.
20. Mauskopf JA, Sullivan SD, Annemans L, Caro J, Mullins CD, Nuijten M, et al. Principles of good practice for budget impact analysis: report of the ISPOR Task Force on good research practices--budget impact analysis. Value Health 2007;10(5):336-347. [doi: 10.1111/j.1524-4733.2007.00187.x] [Medline: 17888098]
21. Sullivan SD, Mauskopf JA, Augustovski F, Jaime Caro J, Lee KM, Minchin M, et al. Budget impact analysis-principles of good practice: report of the ISPOR 2012 Budget Impact Analysis Good Practice II Task Force. Value Health 2014;17(1):5-14 [FREE Full text] [doi: 10.1016/j.jval.2013.08.2291] [Medline: 24438712]

22. Weinstein MC, Russell LB, Gold MR, Siegel JE. Cost-Effectiveness in Health and Medicine. New York, NY: Oxford University Press; 1996.
23. French DD, Dixon BE, Perkins SM, Myers LJ, Weiner M, Zillich AJ, et al. Short-term medical costs of a VHA health information exchange: a CHEERS-compliant article. Medicine (Baltimore) 2016 Jan;95(2):e2481 [FREE Full text] [doi: 10.1097/MD.0000000000002481] [Medline: 26765453]
24. Buntin MB, Zaslavsky AM. Too much ado about two-part models and transformation? Comparing methods of modeling Medicare expenditures. J Health Econ 2004 May;23(3):525-542. [doi: 10.1016/j.jhealeco.2003.10.005] [Medline: 15120469]
25. Manning WG, Mullahy J. Estimating log models: to transform or not to transform? J Health Econ 2001 Jul;20(4):461-494. [Medline: 11469231]
26. French DD, LaMantia MA, Livin LR, Herceg D, Alder CA, Boustani MA. Healthy Aging Brain Center improved care coordination and produced net savings. Health Aff (Millwood) 2014 Apr;33(4):613-618. [doi: 10.1377/hlthaff.2013.1221] [Medline: 24711322]
27. Miles M, Huberman A, Saldana J. Qualitative Data Analysis: A Methods Sourcebook. Thousand Oaks, CA: SAGE Publications; 2013.
28. Committee on Quality of Health Care in America, Institute of Medicine. Crossing the Quality Chasm: A New Health System for the 21st Century. Washington, DC: The National Academies Press; 2001.

Abbreviations

MAIT: Mutual Accountability and Information Therapy
POS: point-of-service

Edited by G Eysenbach; submitted 10.06.19; peer-reviewed by D Bradham, D Nault, J Lander; comments to author 01.07.19; revised version received 26.08.19; accepted 30.08.19; published 17.10.19.

Please cite as:
Greene JC, Haun JN, French DD, Chambers SL, Roswell RH
Reduced Hospitalizations, Emergency Room Visits, and Costs Associated with a Web-Based Health Literacy, Aligned-Incentive Intervention: Mixed Methods Study
J Med Internet Res 2019;21(10):e14772
URL: https://www.jmir.org/2019/10/e14772
doi:10.2196/14772
PMID:31625948

Original Paper

Electronic Health Record–Based Strategy to Promote Medication Adherence Among Patients With Diabetes: Longitudinal Observational Study

Stacy Cooper Bailey[1], MPH, PhD; Amisha Wallia[2], MS, MD; Sarah Wright[1], MPH; Guisselle A Wismer[1], MPH; Alexandra C Infanzon[3], MPH; Laura M Curtis[1], MS; Samantha A Brokenshire[3], PharmD; Arlene E Chung[4,5], MMCi, MHA, MD; Daniel S Reuland[4], MPH, MD; Allison J Hahr[2], MD; Kenneth Hornbuckle[6], DVM, PhD; Karen Lockwood[6], BSN, MPM; Lori Hall[6], BSN; Michael S Wolf[1], MPH, PhD

[1]Division of General Internal Medicine and Geriatrics, Feinberg School of Medicine, Northwestern University, Chicago, IL, United States

[2]Division of Endocrinology, Metabolism, and Molecular Medicine, Department of Medicine, Feinberg School of Medicine, Northwestern University, Chicago, IL, United States

[3]Division of Pharmaceutical Outcomes and Policy, UNC Eshelman School of Pharmacy, University of North Carolina, Chapel Hill, NC, United States

[4]Division of General Internal Medicine & Clinical Epidemiology, Department of Medicine, University of North Carolina School of Medicine, Chapel Hill, NC, United States

[5]Program on Health & Clinical Informatics, University of North Carolina School of Medicine, Chapel Hill, NC, United States

[6]Eli Lilly and Company, Indianapolis, IN, United States

Corresponding Author:
Stacy Cooper Bailey, MPH, PhD
Division of General Internal Medicine and Geriatrics
Feinberg School of Medicine
Northwestern University
10th floor
750 N Lake Shore Dr
Chicago, IL, 60611
United States
Phone: 1 312 503 5595
Email: stacy-bailey@northwestern.edu

Abstract

Background: Poor medication adherence is common; however, few mechanisms exist in clinical practice to monitor how patients take medications in outpatient settings.

Objective: This study aimed to pilot test the Electronic Medication Complete Communication (EMC^2) strategy, a low-cost, sustainable approach that uses functionalities within the electronic health record to promote outpatient medication adherence and safety.

Methods: The EMC^2 strategy was implemented in 2 academic practices for 14 higher-risk diabetes medications. The strategy included: (1) clinical decision support alerts to prompt provider counseling on medication risks, (2) low-literacy medication summaries for patients, (3) a portal-based questionnaire to monitor outpatient medication use, and (4) clinical outreach for identified concerns. We recruited adult patients with diabetes who were prescribed a higher-risk diabetes medication. Participants completed baseline and 2-week interviews to assess receipt of, and satisfaction with, intervention components.

Results: A total of 100 patients were enrolled; 90 completed the 2-week interview. Patients were racially diverse, 30.0% (30/100) had a high school education or less, and 40.0% (40/100) had limited literacy skills. About a quarter (28/100) did not have a portal account; socioeconomic disparities were noted in account ownership by income and education. Among patients with a portal account, 58% (42/72) completed the questionnaire; 21 of the 42 patients reported concerns warranting clinical follow-up. Of these, 17 were contacted by the clinic or had their issue resolved within 24 hours. Most patients (33/38, 89%) who completed the portal questionnaire and follow-up interview reported high levels of satisfaction (score of 8 or greater on a scale of 1-10).

Conclusions: Findings suggest that the EMC^2 strategy can be reliably implemented and delivered to patients, with high levels of satisfaction. Disparities in portal use may restrict intervention reach. Although the EMC^2 strategy can be implemented with minimal impact on clinic workflow, future trials are needed to evaluate its effectiveness to promote adherence and safety.

(J Med Internet Res 2019;21(10):e13499) doi:10.2196/13499

KEYWORDS

health literacy; medication adherence; patient portal; clinical decision support

Introduction

Background

Medication nonadherence has long been recognized as a major public health and patient safety concern, costing the US health care system billions annually and compromising the benefit and risk profile of patients' treatments [1,2]. Nonadherence has been linked to negative health outcomes across a number of chronic conditions, including poorer disease control, increased hospitalizations, and greater morbidity and mortality [3-6]. However, decades of research has shown that poor medication adherence in its many forms—failure to fill, early abandonment of therapy, and discontinuation—continues to be common and pervasive [3,7-9].

Despite its prevalence and consequences, medication adherence is often inadequately addressed during clinical encounters. A number of studies have shown that provider-patient communication on medication use is suboptimal [10-13]. As a result, patients are often poorly informed about treatment plans and lack the knowledge necessary to manage medications safely and effectively on their own [9,14]. From the provider perspective, it is difficult to objectively, efficiently, and accurately collect information on patients' medication adherence in time-constrained clinical settings [13]. Information that is obtained may not be clinically actionable, and few programs have been widely implemented to monitor outpatient medication use and intervene when necessary should safety or adherence concerns be detected [9].

Technology-Based Solutions

Recently, increased attention has been placed on how health information technology can be used to collect data on patient medication use with the goal of improving outpatient medication adherence and safety [9,15-17]. The Health Information Technology for Economic and Clinical Health Act and the Office of the National Coordinator's Meaningful Use initiative led to a dramatic increase in the adoption of electronic health records (EHRs) in health care systems nationwide [18]. In the context of outpatient medication use, EHRs can provide a unique opportunity to improve medication monitoring and safety, particularly via their associated tethered patient portals, which can better connect patients to their providers in outpatient settings. This is a distinct advantage of EHR and tethered portal-based interventions in comparison with other consumer technologies (eg, short message service (SMS) text messaging and mobile apps), which are often not as easily linked and integrated with clinical care workflows or clinical data.

Although EHRs and tethered portals have tremendous potential, there have been a limited number of methodologically rigorous studies that have sought to develop and evaluate EHR-based interventions to promote and monitor safe outpatient medication use. Graumlich et al evaluated MedTable, a patient education tool embedded within the EHR, which promoted medication review and reconciliation between patients and providers in a randomized controlled trial among 674 patients with type 2 diabetes [19]. They found that intervention patients had greater knowledge of medication indications but did not experience improved medication adherence or hemoglobin A_{1c}. Similarly, both Weingart et al and Schnipper et al developed and evaluated EHR modules that promoted medication review and reconciliation by patients and providers in 2 randomized controlled trials [20,21]. Although Weingart et al found that intervention patients in their trial had more accurate medication lists with fewer discrepancies with the potential for severe harm, neither study reported a reduction in adverse drug events among intervention patients [20,21].

In terms of medication adherence, Volmer et al found that a multifaceted intervention, which included an EHR-based component, significantly increased adherence to statins and angiotensin-converting enzyme inhibitors/angiotensin receptor blockers among intervention patients compared with usual care; however, their strategy utilized automated telephone reminders, personalized letters, and live outreach, not solely EHR functions [16]. Similarly, Simon et al investigated the effectiveness of a depression care management program delivered via messaging through an EHR [22]. They found that intervention patients had higher antidepressant adherence rates and fewer depression symptoms at 5 months than usual care patients. However, this trial was conducted in a large integrated health care system with greater care coordination between primary care and mental health providers; authors acknowledged that the intervention strategy may not be feasible in other practice settings.

Study Purpose

Although EHRs and their associated portals offer a unique opportunity to bridge patients to their providers, prior studies have reported variable success in using EHR-based strategies to support and monitor medication use in outpatient settings [16,19-22]. Those interventions that appear to have been the most successful to date may be less feasible for widespread implementation. To address this concern, our team recently devised and pilot-tested the Electronic Medication Complete Communication (EMC^2) strategy, a simple, sustainable approach that uses EHR technology to promote medication adherence and safety by providing routine assessment and monitoring of patients' medication use at home. Herein, we describe the core

components of the EMC^2 strategy and results from our initial pilot test, which focused on building the technological infrastructure within the EHR and portal to support the intervention as well as determining the feasibility of the strategy and its acceptability among a diverse set of patients from both primary care and specialty practices.

Methods

Overview

The EMC^2 strategy is a multifaceted intervention that consists of several components designed to (1) promote provider counseling on medication adherence and safe use, (2) provide patient-friendly education at the point of prescribing, (3) monitor patient medication use in outpatient settings, and (4) prompt clinic follow-up with patients reporting medication-related concerns. We used the Institute of Medicine's *Medication Use Process Model* to inform our strategy and the above components; this model deconstructs common system failures in the processes of prescribing, dispensing, self-administering, and monitoring medications that commonly lead to medication errors, adverse drug events, and suboptimal adherence and treatment [23].

The EMC^2 strategy is intended to be flexible, with intervention components that can be adapted for specific settings and patient populations. Concurrent to this project, other iterations of the EMC^2 strategy have been developed to provide monitoring via interactive voice recognition technology among low-income, primary care patients and to monitor opioid use among emergency department patients [24,25]. For this study, the EMC^2 strategy was delivered via Epic EHR Software (Epic Systems Corporation) at both study sides. The functionalities described are not specific to Epic, however, and could be deployed in other EHRs with tethered patient portals. Our team has implemented EHR-based interventions using similar functionalities across a number of distinct EHR platforms [25,26].

Study Medications

For the purpose of this study, we tailored the EMC^2 strategy to support medication adherence and safe use among patients with diabetes, a chronic condition that is becoming increasingly prevalent and for which medication use is a cornerstone of self-management [3,27]. We selected various prescription medications used to treat diabetes and their associated generic formulations as targets for the intervention. These medications were selected as most are considered higher risk according to the Food and Drug Administration (FDA) and consequently require an FDA-approved Medication Guide. Despite these regulations, patients may not receive adequate counseling or monitoring on higher-risk medications during and after the clinical encounter [10,23].

Intervention Components

The EMC^2 strategy tested in this study included 4 core components, as described below.

Best Practice Alert for Physicians

Best practice alerts (BPAs) are clinical decision support alerts within Epic EHR systems that provide a mechanism to show relevant data in a pop-up window at the point-of-care. When a prescriber placed a new order or refill for 1 of the study medications, a BPA fired to prompt the prescriber to counsel the patient on medication use and key medication risks. The BPA also listed key adverse events associated with the medication according to the FDA Medication Guide and/or prescriber insert and included a hyperlink within the BPA window to additional medication information.

Automated Delivery of Food and Drug Administration Medication Guide + Medication Summary

Upon placing an order for a higher-risk medication, an FDA Medication Guide and a 1-page, patient-friendly *Medication Summary* were automatically queued for printing along with the patients' after-visit summary (AVS). The AVS is a handout that provides the patient with their care plans, medication lists, and other information that is provided at the conclusion of the clinical encounter. The Medication Summary provided an overview of medication instructions, risks, and benefits in patient-friendly language. Medication Summaries were previously developed by our research team and have been shown to support patient comprehension of medication information [28].

Follow-Up Portal Assessment

Approximately 1 week after their clinic visit, patients with a portal account were sent an email prompting them to log on to the portal to complete and submit a brief, voluntary questionnaire. This questionnaire consisted of 9 questions, which were developed by our study team, to assess common root causes identified in the scientific literature as being related to medication adherence and safety (ie, forgetfulness, cost, side effects; see Multimedia Appendix 1) [29]. The online survey was only available by invitation to those who had enrolled in the study and had a portal account. All items on the survey appeared on 1 page and could be easily completed by patients. Patients were able to review responses to their questions before submission but were only allowed to complete the questionnaire once. Patients without a portal account did not receive any further contact.

Clinic Follow-Up

The results of the patient portal questionnaire were automatically sent via an EHR message to clinic staff. Clinic staff reviewed results then followed up with any identified medication-related concerns. Clinics were asked to determine their own specific protocols for follow-up; these were separate from research activities.

Study Design

To determine the feasibility and acceptability of the EMC^2 approach, a single-arm study was conducted to build and pilot-test the EMC^2 strategy in 2 academic health care settings: an endocrinology clinic in Chicago, Illinois, and a general internal medicine clinic in Chapel Hill, North Carolina. Both clinics utilized Epic EHR systems.

The EHR build included programming the BPAs, Medication Summaries, and FDA Medication Guides to trigger and launch whenever a medication order was placed for 1 of the study medications. This functionality was delivered to all patients who visited a study clinic and received a prescription for 1 of the study medications, regardless of their study involvement; this ensured that patients received all in-clinic EMC^2 components on the day of their index clinic visit. Only those patients who consented to participate in the study following their clinic appointment were eligible to receive the remaining EMC^2 strategy components (ie, portal assessment and clinic follow-up) and to participate in evaluation activities. The institutional review boards of Northwestern University and the University of North Carolina at Chapel Hill approved study procedures.

Participants

Patients were considered eligible for the study if they (1) were aged 18 years or older, (2) spoke English, (3) were primarily responsible for administering their own medication, (4) received a new or refill prescription for a study medication, and (5) had access to the internet. Adults with severe, uncorrectable vision, hearing, or cognitive impairments, or who were unable to provide informed consent, were excluded from participating. Adults without a portal account were allowed to enroll in the study and were told that they could sign up for an account if they desired.

To recruit patients for the evaluation of the EMC^2 strategy, EHRs were configured with eligibility criteria such that a message was sent to research assistants whenever a study medication was ordered for a patient in the study clinic. Research assistants then contacted identified patients by either approaching them following their clinic visit or calling them via telephone following their appointment. Research assistants then introduced the study, verified eligibility, and enrolled interested and eligible patients in the study. Participants completed a baseline, in-person interview then a telephone interview 2 weeks post clinic visit. Data were collected using REDCap software (Research Electronic Data Capture) hosted by Northwestern University [30]. Participants were compensated for participating in study interviews; the same incentive was received whether participants completed the follow-up portal assessment or not.

Measurement

Participant Characteristics

Patient sociodemographic variables (eg, race, income, age, and sex) were collected during the baseline interview. Patient health literacy skills were measured via the Newest Vital Sign (NVS), a commonly used assessment that asks patients to interpret information presented on a standard Nutrition Facts label [31]. Information regarding patients' prior use of technology and the patient portal was also collected based on patient self-report.

Process Outcomes

Outcomes related to the fidelity of the intervention were collected via patient self-report and EHR data. Specifically, during the baseline interview, patients were asked to self-report receipt of provider counseling during their index clinic visit (yes/no) and receipt of a Medication Summary following their clinic visit (yes/no). Data from the EHR, collected after the 2-week interview, were used to determine if patients received and completed the portal questionnaire and to determine if any clinic follow-up occurred based upon responses to the portal survey.

Patient Satisfaction

Patient satisfaction with the various EMC^2 strategy components was also assessed. During the baseline interview, patients were asked about communication with their providers on medication use via a modified version of the supplemental Health Literacy items from the Consumer Assessment of Healthcare Providers and Systems questionnaire [18]. They were also asked to rate their satisfaction with the Medication Summary and their perceived helpfulness of the tool on a scale of 1 to 10. During the 2-week interview, patients were asked about their continued satisfaction with the Medication Summary and whether it was still in use (yes/no). Finally, during the 2-week interview, patients reported on their experiences with the portal questionnaire, including any barriers to its completion (eg, difficulty logging on to the portal and challenges understanding or answering questions), and their overall satisfaction with the tool (scale of 1-10).

Statistical Analysis

Descriptive statistics were calculated for patient sociodemographic variables, process outcomes, and variables related to patient satisfaction with the intervention. To assess whether there were any systematic, statistically significant differences between patients who had access to the portal and those who did not, we used Pearson chi-squared tests or Fisher exact tests for categorical variables and Student *t* tests for age. The same tests were used to compare the differences between the patients who completed the portal questionnaire versus those who did not in those with portal access. Specifically, we examined if there were variations by age, sex, race, education, income, health literacy skills, and study site. Statistical significance was defined as alpha<.05.

Results

Overview

Recruitment began in December 2016 and concluded in April 2017. A total of 251 patients were approached or contacted by a research assistant; 66 patients declined to participate, and 31 patients were ineligible. A total of 100 patients completed the baseline interview (n=43 in North Carolina, n=57 in Illinois), and 90 patients completed the 2-week follow-up interview. There were no significant differences in sociodemographic characteristics between those who completed the follow-up interview and those who did not.

Participant Characteristics

Table 1 describes the characteristics of the baseline study sample (N=100). The mean age was 56.2 (SD 11.20; range 24-82) years. More than half (57/100, 57.0%) of the sample was female, about one-third were African American (38/100, 38.0%), and 30.0%

(30/100) reported a high school or less level of education. A total of 40 patients (40/100) had limited health literacy skills (low or marginal) according to the NVS. About a quarter (28/100, 28.0%) of patients reported not having a portal account. Among those who reported having and using a portal account (n=67), 54% (36/67) stated that they used the portal once per month or less. Disparities were noted in portal account ownership by level of education, household income, race, and study site (Table 1).

Process Outcomes

Almost two-thirds (63/100, 63.0%) of patients reported that their physician counseled them on how to take their medication, and 54.0% of patients (54/100) stated that their physician counseled them on possible side effects of the medication. The majority of patients (72/100, 72.0%) reported receiving a Medication Summary after their index clinic visit. Of those who completed the 2-week follow-up interview, 74% (67/90) reported that they still had the Medication Summary in their possession.

Table 1. Participant characteristics, portal account ownership, and questionnaire completion.

Participant Characteristic	All participants (n=100)	Portal account		Portal questionnaire	
		Proportion with account (n=100)	*P* value	Proportion completed (n=72)	*P* value
Overall, n (%)	—[a]	72 (72.0)	—	42 (58)	—
Age (years)	—	—	.09	—	.50
<55	42 (42.0)	35 (83.3)	—	22 (63)	—
55-64	34 (34.0)	21 (61.8)	—	10 (48)	—
>65	24 (24.0)	16 (66.7)	—	10 (63)	—
Sex	—	—	.36	—	.28
Male	43 (43.0)	33 (76.7)	—	17 (52)	—
Female	57 (57.0)	39 (68.4)		25 (64)	
Race[b]	—	—	<.01	—	.70
Black	38 (38.0)	20 (52.6)	—	10 (50)	—
White	48 (48.0)	39 (81.3)	—	24 (62)	—
Other	13 (13.0)	12 (92.3)	—	7 (58)	—
Education	—	—	.02	—	.01
High school or less	30 (30.0)	16 (53.3)	—	4 (25)	—
Some college	33 (33.0)	25 (75.8)	—	16 (64)	—
College graduate	37 (37.0)	31 (83.8)	—	22 (71)	—
Income[c] **(US $)**	—	—	.03	—	.07
<30,000	24 (24.0)	13 (54.2)	—	4 (31)	—
30-49,999	20 (20.0)	15 (75.0)	—	9 (60)	—
50-99,999	27 (27.0)	21 (77.8)	—	13 (62)	—
>100,000	23 (23.0)	21 (91.3)	—	16 (76)	—
Literacy	—	—	.10	—	.12
Low	12 (12.0)	6 (50.0)	—	1 (17)	—
Marginal	28 (28.0)	19 (67.9)	—	11 (58)	—
Adequate	60 (60.0)	47 (78.3)	—	30 (64)	—
Site	—	—	<.01	—	.09
Chicago	57 (57.0)	51 (89.5)	—	33 (65)	—
North Carolina	43 (43.0)	21 (48.8)	—	9 (43)	—

[a]Not applicable.

[b]n=1 refused race.

[c]n=6 do not know or refused income.

Among patients with a portal account, 58% (42/72) completed the portal questionnaire. Patients with lower levels of education were less likely to complete the questionnaire (*P*=.01). Among those patients who submitted a questionnaire (n=42), 12 did not

report any medication-related concerns on the survey, 10 reported minor concerns that did not warrant follow-up according to individualized clinic protocols, and 21 identified issues that triggered a clinical response. Of the latter, 81% (17/21) were contacted by the clinic or had their medication-related issue resolved within 24 hours. The remaining 4 patients received clinic outreach within 5 days. The most frequent problems identified in the portal questionnaire were not currently having the medication in possession (n=34), difficulties paying for the medication (n=5), and worries about side effects (n=17).

Patient Satisfaction

Overall, patients reported high levels of satisfaction with communication with their providers on medication use and with the intervention components. Among patients who reported receiving counseling on how to take their medications, 97% (61/63) said the guidance was very or mostly easy to understand. Similarly, among those adults reporting receiving counseling on side effects (n=54), all said the counseling was very or mostly easy to understand.

Almost all (93/96, 97%) of the patients felt the Medication Summaries were very easy or mostly easy to understand. On a scale from 1 to 10 (with 10 being the best), patients scored the Medication Summary an average of 9.0 (SD 1.6) in being clear and 9.2 (SD 1.3) in being helpful. All of the 38 patients who completed both the portal questionnaire and the follow-up interview reported that the portal-based questionnaire was very easy or somewhat easy to access. A total of 89% (33/38) of patients completing the portal questionnaire and the follow-up interview reported high levels of satisfaction with the portal experience (8 or more on a scale of 1-10).

Discussion

Principal Findings

Results from this study indicate that the EMC^2 strategy can be reliably implemented and delivered to patients. Medication Summaries were received by the majority of patients at their index clinic visit, and most patients still had the materials in their possession 2 weeks after their clinic visit. Most patients who had a patient portal account completed the online questionnaire, and clinic staff were able to resolve issues identified in the questionnaires in a timely manner. Importantly, patients reported high levels of satisfaction with intervention components, and the EMC^2 strategy appears to have been successfully implemented in 2 diverse practices with minimal impact on clinic workflow.

Although there were many positive results from this feasibility study, we also uncovered some shortcomings to the EMC^2 strategy. More than a quarter of patients enrolled in the study did not have access to the patient portal and were, therefore, unable to complete the questionnaire. Significant socioeconomic disparities were found between those patients who had portal access and those who did not. These findings are consistent with prior studies and likely mirror socioeconomic disparities in the internet and portal access and use [32,33]. Similarly, we also found disparities in questionnaire completion among patients who had portal access. Specifically, patients with less educational attainment were less likely to complete the portal questionnaire, even though they had portal access. It is possible that patients with lower education levels may need additional technical assistance and support to effectively utilize the patient portal and to complete portal-based questionnaires. It is possible that follow-up assessments may need to be conducted via telephone or another modality for a limited number of patients, with results recorded in the EHR for the care team to review and address.

Although these disparities limit the current reach of the EMC^2 strategy, data indicates that internet access and use is on the rise and that US adults are becoming increasingly familiar with using technology to support their health [34,35]. Thus, although the EMC^2 strategy may not work for all patients at this point in time, it is plausible that its reach will increase greatly in years to come. This is also likely given the large investments made in health information technology by health care systems nationwide and the increasing attention placed on promoting the use of patient portals by health systems.

Limitations

There are limitations to this study that should be noted. Specifically, this was a small pilot study that was focused on building the EMC^2 strategy components in the EHR and portal and determining the feasibility of implementing the strategy in 2 diverse clinics within 2 separate health systems. As such, our focus was on understanding the delivery of the EMC^2 strategy and patient satisfaction with the intervention components; additional studies are needed to fully evaluate the effectiveness of the strategy itself. Although participants were recruited from both specialty and primary care practices in 2 diverse locations, we only included English-speaking patients in our study, which limits study generalizability. Future work is needed to develop intervention materials and to test the strategy among non-English speaking populations.

Conclusions

Physician time for counseling and monitoring medication use in outpatient settings is extremely limited. Studies have shown that providers would need to spend approximately 18 hours per working day to deliver counseling and care consistent with US Preventive Services Task Force and chronic conditions care management recommendations [36,37]. To improve quality of communication around medication adherence and safety, hard-wired approaches such as the EMC^2 strategy are, therefore, clearly needed. EMC^2 tools could help streamline physician counseling, and written materials could support patient learning on safe and appropriate medication use outside of the clinic visit. Utilizing the patient portal to provide outpatient monitoring of medication use can help identify those patients in greatest need of additional follow-up. Importantly, the strategy can also help patients engage with other members of the care team (ie, nurses and care coordinators) so that counseling is not entirely dependent upon physicians during time-limited visits. Additional research is currently underway utilizing a rigorous, randomized

study design to formally test the effectiveness of the EMC[2] strategy against usual care.

Acknowledgments

This study was funded by Eli Lilly and Company. Authors include employees of Eli Lilly and Company.

Authors' Contributions

Eli Lilly reviewed and approved the manuscript for submission. SCB, AW, SW, GAW, ACI, AEC, DSR, AJH, KH, KL, LH, and MSW all contributed to the conception and design of the study and the interpretation of findings. ACI, SAB, GAW, SW, LMC, MSW, and SCB assisted in the acquisition, analysis, and interpretation of data. AJH, AW, AEC, and DSR served as clinical liaisons for the study sites and provided substantive, clinical expertise. All authors were involved in the drafting of this manuscript and gave final approval for submission.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Portal questionnaire.
[PDF File (Adobe PDF File), 25 KB - jmir_v21i9e13499_app1.pdf]

References

1. Cutler RL, Fernandez-Llimos F, Frommer M, Benrimoj C, Garcia-Cardenas V. Economic impact of medication non-adherence by disease groups: a systematic review. BMJ Open 2018 Jan 21;8(1):e016982 [FREE Full text] [doi: 10.1136/bmjopen-2017-016982] [Medline: 29358417]
2. Lloyd JT, Maresh S, Powers CA, Shrank WH, Alley DE. How much does medication nonadherence cost the medicare fee-for-service program? Med Care 2019 Mar;57(3):218-224. [doi: 10.1097/MLR.0000000000001067] [Medline: 30676355]
3. Capoccia K, Odegard PS, Letassy N. Medication adherence with diabetes medication: a systematic review of the literature. Diabetes Educ 2016 Feb;42(1):34-71. [doi: 10.1177/0145721715619038] [Medline: 26637240]
4. Gourzoulidis G, Kourlaba G, Stafylas P, Giamouzis G, Parissis J, Maniadakis N. Association between copayment, medication adherence and outcomes in the management of patients with diabetes and heart failure. Health Policy 2017 Apr;121(4):363-377. [doi: 10.1016/j.healthpol.2017.02.008] [Medline: 28314467]
5. Simpson SH, Lin M, Eurich DT. Medication adherence affects risk of new diabetes complications: a cohort study. Ann Pharmacother 2016 Sep;50(9):741-746. [doi: 10.1177/1060028016653609] [Medline: 27307411]
6. Du L, Cheng Z, Zhang Y, Li Y, Mei D. The impact of medication adherence on clinical outcomes of coronary artery disease: a meta-analysis. Eur J Prev Cardiol 2017 Jun;24(9):962-970. [doi: 10.1177/2047487317695628] [Medline: 28436725]
7. Iglay K, Cartier SE, Rosen VM, Zarotsky V, Rajpathak SN, Radican L, et al. Meta-analysis of studies examining medication adherence, persistence, and discontinuation of oral antihyperglycemic agents in type 2 diabetes. Curr Med Res Opin 2015;31(7):1283-1296. [doi: 10.1185/03007995.2015.1053048] [Medline: 26023805]
8. Vermeire E, Hearnshaw H, van Royen P, Denekens J. Patient adherence to treatment: three decades of research. A comprehensive review. J Clin Pharm Ther 2001 Oct;26(5):331-342. [doi: 10.1046/j.1365-2710.2001.00363.x] [Medline: 11679023]
9. Giardini A, Martin MT, Cahir C, Lehane E, Menditto E, Strano M, et al. Toward appropriate criteria in medication adherence assessment in older persons: position paper. Aging Clin Exp Res 2016 Jun;28(3):371-381. [doi: 10.1007/s40520-015-0435-z] [Medline: 26630945]
10. Tarn DM, Wenger A, Good JS, Hoffing M, Scherger JE, Wenger NS. Do physicians communicate the adverse effects of medications that older patients want to hear? Drugs Ther Perspect 2015 Feb;31(2):68-76 [FREE Full text] [doi: 10.1007/s40267-014-0176-7] [Medline: 25750513]
11. Tarn DM, Mattimore TJ, Bell DS, Kravitz RL, Wenger NS. Provider views about responsibility for medication adherence and content of physician-older patient discussions. J Am Geriatr Soc 2012 Jun;60(6):1019-1026 [FREE Full text] [doi: 10.1111/j.1532-5415.2012.03969.x] [Medline: 22646818]
12. Institute of Medicine, Board on Population Health and Public Health Practice, Roundtable on Health Literacy. In: Hernandez LM, editor. Standardizing Medication Labels: Confusing Patients Less: Workshop Summary. Washington, DC: National Academies Press; 2008.
13. Hines R, Stone NJ. Patients and physicians beliefs and practices regarding adherence to cardiovascular medication. JAMA Cardiol 2016 Jul 1;1(4):470-473. [doi: 10.1001/jamacardio.2016.0634] [Medline: 27438324]

14. National Academies of Sciences Engineering and Medicine, Health and Medicine Division, Board on Population Health and Public Health Practice, Roundtable on Health Literacy. In: Alper J, editor. Communicating Clearly about Medicines: Proceedings of a Workshop. Washington, DC: National Academies Press; 2017.
15. Granger BB, Bosworth HB. Medication adherence: emerging use of technology. Curr Opin Cardiol 2011 Jul;26(4):279-287 [FREE Full text] [doi: 10.1097/HCO.0b013e328347c150] [Medline: 21597368]
16. Vollmer WM, Owen-Smith AA, Tom JO, Laws R, Ditmer DG, Smith DH, et al. Improving adherence to cardiovascular disease medications with information technology. Am J Manag Care 2014 Nov;20(11 Spec No 17):SP502-SP510 [FREE Full text] [Medline: 25811824]
17. Perri-Moore S, Kapsandoy S, Doyon K, Hill B, Archer M, Shane-McWhorter L, et al. Automated alerts and reminders targeting patients: a review of the literature. Patient Educ Couns 2016 Jun;99(6):953-959 [FREE Full text] [doi: 10.1016/j.pec.2015.12.010] [Medline: 26749357]
18. Centers for Disease Control and Prevention. 2017. Meaningful Use: Introduction URL: http://www.webcitation.org/78VykKVRq [accessed 2019-05-18] [WebCite Cache ID 78VykKVRq]
19. Graumlich JF, Wang H, Madison A, Wolf MS, Kaiser D, Dahal K, et al. Effects of a patient-provider, collaborative, medication-planning tool: a randomized, controlled trial. J Diabetes Res 2016;2016:2129838 [FREE Full text] [doi: 10.1155/2016/2129838] [Medline: 27699179]
20. Weingart SN, Carbo A, Tess A, Chiappetta L, Tutkus S, Morway L, et al. Using a patient internet portal to prevent adverse drug events: a randomized, controlled trial. J Patient Saf 2013 Sep;9(3):169-175. [doi: 10.1097/PTS.0b013e31829e4b95] [Medline: 23965840]
21. Schnipper JL, Gandhi TK, Wald JS, Grant RW, Poon EG, Volk LA, et al. Effects of an online personal health record on medication accuracy and safety: a cluster-randomized trial. J Am Med Inform Assoc 2012;19(5):728-734 [FREE Full text] [doi: 10.1136/amiajnl-2011-000723] [Medline: 22556186]
22. Simon GE, Ralston JD, Savarino J, Pabiniak C, Wentzel C, Operskalski BH. Randomized trial of depression follow-up care by online messaging. J Gen Intern Med 2011 Jul;26(7):698-704 [FREE Full text] [doi: 10.1007/s11606-011-1679-8] [Medline: 21384219]
23. Institute of Medicine, Board on Health Care Services, Committee on Identifying and Preventing Medication Errors. In: Cronenwett LR, Bootman JL, Wolcott J, Aspden P, editors. Preventing Medication Errors. Washington, DC: National Academies Press; 2007.
24. McCarthy DM, Courtney DM, Lank PM, Cameron KA, Russell AM, Curtis LM, et al. Electronic medication complete communication strategy for opioid prescriptions in the emergency department: rationale and design for a three-arm provider randomized trial. Contemp Clin Trials 2017 Aug;59:22-29. [doi: 10.1016/j.cct.2017.05.003] [Medline: 28479220]
25. Bailey SC, Paasche-Orlow MK, Adams WG, Brokenshire SA, Hedlund LA, Hickson RP, et al. Corrigendum to 'the electronic medication complete communication (EMC2) study: rationale and methods for a randomized controlled trial of a strategy to promote medication safety in ambulatory care'. Contemp Clin Trials 2017 Jun;57:99. [doi: 10.1016/j.cct.2017.03.002] [Medline: 28325511]
26. Persell SD, Eder M, Friesema E, Connor C, Rademaker A, French DD, et al. EHR-based medication support and nurse-led medication therapy management: rationale and design for a three-arm clinic randomized trial. J Am Heart Assoc 2013 Oct 24;2(5):e000311 [FREE Full text] [doi: 10.1161/JAHA.113.000311] [Medline: 24157649]
27. Rowley WR, Bezold C, Arikan Y, Byrne E, Krohe S. Diabetes 2030: insights from yesterday, today, and future trends. Popul Health Manag 2017 Feb;20(1):6-12 [FREE Full text] [doi: 10.1089/pop.2015.0181] [Medline: 27124621]
28. Wolf MS, Bailey SC, Serper M, Smith M, Davis TC, Russell AL, et al. Comparative effectiveness of patient-centered strategies to improve FDA medication guides. Med Care 2014 Sep;52(9):781-789. [doi: 10.1097/MLR.0000000000000182] [Medline: 25119953]
29. Hugtenburg JG, Timmers L, Elders PJ, Vervloet M, van Dijk L. Definitions, variants, and causes of nonadherence with medication: a challenge for tailored interventions. Patient Prefer Adherence 2013;7:675-682 [FREE Full text] [doi: 10.2147/PPA.S29549] [Medline: 23874088]
30. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009 Apr;42(2):377-381 [FREE Full text] [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
31. Weiss BD, Mays MZ, Martz W, Castro KM, DeWalt DA, Pignone MP, et al. Quick assessment of literacy in primary care: the newest vital sign. Ann Fam Med 2005;3(6):514-522 [FREE Full text] [doi: 10.1370/afm.405] [Medline: 16338915]
32. Bailey SC, O'Conor R, Bojarski EA, Mullen R, Patzer RE, Vicencio D, et al. Literacy disparities in patient access and health-related use of internet and mobile technologies. Health Expect 2015 Dec;18(6):3079-3087 [FREE Full text] [doi: 10.1111/hex.12294] [Medline: 25363660]
33. Smith SG, O'Conor R, Aitken W, Curtis LM, Wolf MS, Goel MS. Disparities in registration and use of an online patient portal among older adults: findings from the LitCog cohort. J Am Med Inform Assoc 2015 Jul;22(4):888-895 [FREE Full text] [doi: 10.1093/jamia/ocv025] [Medline: 25914099]

34. Anderson M, Perrin A. Pew Research Center. 2017. Tech Adoption Climbs Among Older Adults URL: https://www.pewinternet.org/2017/05/17/tech-adoption-climbs-among-older-adults/ [accessed 2019-05-15] [WebCite Cache ID 78OVsDJZP]
35. Pew Research Center. 2019. Mobile Fact Sheet URL: https://www.pewinternet.org/fact-sheet/mobile/ [accessed 2019-08-08] [WebCite Cache ID 78OW1JDAf]
36. Yarnall KS, Pollak KI, Østbye T, Krause KM, Michener JL. Primary care: is there enough time for prevention? Am J Public Health 2003 Apr;93(4):635-641. [doi: 10.2105/ajph.93.4.635] [Medline: 12660210]
37. Østbye T, Yarnall KS, Krause KM, Pollak KI, Gradison M, Michener JL. Is there time for management of patients with chronic diseases in primary care? Ann Fam Med 2005;3(3):209-214 [FREE Full text] [doi: 10.1370/afm.310] [Medline: 15928223]

Abbreviations

AVS: after-visit summary
BPA: best practice alert
EHR: electronic health record
EMC2: electronic medication complete communication
FDA: Food and Drug Administration
NVS: Newest Vital Sign

Edited by CL Parra-Calderón; submitted 04.02.19; peer-reviewed by C Alvarez, H Papadopoulos; comments to author 31.03.19; revised version received 29.05.19; accepted 15.06.19; published 21.10.19.

Please cite as:
Bailey SC, Wallia A, Wright S, Wismer GA, Infanzon AC, Curtis LM, Brokenshire SA, Chung AE, Reuland DS, Hahr AJ, Hornbuckle K, Lockwood K, Hall L, Wolf MS
Electronic Health Record–Based Strategy to Promote Medication Adherence Among Patients With Diabetes: Longitudinal Observational Study
J Med Internet Res 2019;21(10):e13499
URL: https://www.jmir.org/2019/10/e13499
doi:10.2196/13499
PMID:31638592

Original Paper

Technology Acceptance and Information System Success of a Mobile Electronic Platform for Nonphysician Clinical Students in Zambia: Prospective, Nonrandomized Intervention Study

Sandra Barteit[1], MA; Florian Neuhann[1], MD; Till Bärnighausen[1,2,3], MSc, MD, ScD, PhD; Annel Bowa[4], MSc; Sigrid Wolter[5], Dr rer nat; Hinson Siabwanta[4], MSc; Albrecht Jahn[1], MSc, MD, PhD

[1]Heidelberg Institute of Global Health, Heidelberg, Germany

[2]Department of Global Health and Population, Harvard T.H. Chan School of Public Health, Boston, MA, United States

[3]Africa Health Research Institute (AHRI), KwaZulu-Natal, South Africa

[4]Lewy Mwanawasa University, Lusaka, Zambia

[5]SolidarMed, Lusaka, Zambia

Corresponding Author:
Sandra Barteit, MA
Heidelberg Institute of Global Health
Im Neuenheimer Feld 130.3
Heidelberg, 69120
Germany
Phone: 49 06221 ext 5634030
Email: barteit@uni-heidelberg.de

Abstract

Background: Zambia is still experiencing a severe shortage of health workers, which is impacting the national health care system. Very few people are trained, educational infrastructure is inadequate, and senior human resources for training are not yet sufficient to produce the number of health care workers needed, especially for currently underserved rural areas. Therefore, to strengthen the medical education program of medical licentiates, we implemented a tablet-based electronic learning platform (e-platform) with a medical decision-support component.

Objective: As the primary objective, this study aimed to explore the acceptance and information system (IS) success of an e-platform focused on offline-based tablet usage for nonphysician clinical students in a low-resource context in Zambia, Africa. Furthermore, we aimed to evaluate student demographic factors and prior technological experience, as well as medical lecturers' acceptance of technology of the e-platform.

Methods: We collected data for the study before and after the intervention. Before the intervention, we collected student demographic data and prior technological experience using a questionnaire. After the intervention, we collected results of the questionnaire on technology acceptance of students and IS success of the e-platform, as well as technology acceptance of medical lecturers. We calculated statistical measures such as means, standard deviations, and correlations of investigated variables. The study report was compiled using the Consolidated Standards of Reporting Trials-Electronic Health checklist.

Results: Overall, questionnaire results of students and medical lecturers indicated acceptance of the e-platform and showed higher ratings for overall net benefits and information quality (students) and perceived ease of use and perceived usefulness (medical lecturers) as compared with ratings of other categories. The lowest scores were conveyed for system use and service quality (students) and attitude and behavioral intention (medical lecturers).

Conclusions: Acceptance of the e-platform as a learning technology for strengthening medical education in a low-resource context in Zambia was generally high for students and medical lecturers, but shortcomings were also identified. Results indicated low overall usage of the e-platform as a learning and teaching tool. One hindering factor was the tablets' overall weak reliability with regard to its service life and battery life span, and another was the teachers' low engagement with the e-platform. Next steps may include other hardware and more technology-based training for medical lecturers. The evaluation results indicated that the e-platform may open new promise for further strengthening and expanding medical education in this context, especially with more affordable and viable technologies that are available.

(J Med Internet Res 2019;21(10):e14748) doi:10.2196/14748

KEYWORDS

computers, handheld; tablets; education, medical; sub-Saharan Africa; Zambia; mHealth; evidence-based practice; medicine

Introduction

Background

Zambia, a country in south-central Africa, faces significant challenges in its overall disease burden and scant numbers of health workers across all sectors and rural regions in particular [1-3]. The current severe shortage of Zambian health workers and their skewed distribution toward urban areas underline the dire need for an upscale of skilled and knowledgeable health workers, especially in rural areas. The ratio of health workers in Zambia compared with the population is 1.2 per 1000 people, which is far from the World Health Organization Sustainable Development Goals index threshold of 4.45 health workers per 1000 people [2,4]. In a country with over 750,000 km^2 area, Zambia is sparsely populated by approximately 20 inhabitants per square kilometer [5], which further aggravates the dense coverage of and proximate access to health care. Zambia's rural areas are severely underserved with only 7 clinicians per 10,000 people (urban areas: 16/10,000 people) and are insufficiently covered by health facilities. The recent Zambian National Health Strategic Plan [1] points out that at times "the situation is so severe that some facilities in rural areas have insignificant numbers of staff and in the worst scenario are managed by unqualified staff." In 2002, to address the need for better-qualified nonphysician clinicians to cover a broader scope of health care than that provided by clinical officers, Zambia started training associate physicians at the Chainama College of Health Sciences (CCHS), Lusaka, to be known as medical licentiate practitioners (MLPs) [6]. MLPs manage advanced and common medical conditions and cover competencies such as surgical care, cesarean sections, and management of complex treatments. Most MLPs work at rural district hospitals or peripheral health centers. Retention of MLPs is high, and over 270 MLPs work in the Zambian health system, thus strengthening health care in rural areas [1]. As a result, the Ministry of Health has set an objective to more than double the MLP workforce to 600 by 2025 [1]. Although the number of MLPs trained at CCHS has increased, especially in rural areas, CCHS as the main national training institution for MLPs has experienced restricted resources and infrastructure to adequately upscale the sorely needed increase in the number of students.

Research Objectives

To strengthen the quantity and quality of training for MLPs, we introduced a blended learning approach that includes a self-directed electronic health platform (e-platform) [7]. The primary objective of the e-platform is to alleviate the shortcomings of the medical licentiate program training caused by shortages of medical lecturers and learning materials in MLP training sites. The e-platform covers 2 components: (1) electronic learning (e-learning) for medical education with Web-based and offline static and interactive learning materials including lecture presentations, books, virtual patient cases, pictures, and videos and (2) health care practice support with medical treatment guidelines and algorithms to diagnose and treat patients. Details on e-platform contents are described elsewhere [8,9]. MLP students were provided with 7-inch tablets preloaded with offline contents of the e-platform through the Moodle mobile app [10]. The blended learning approach was piloted from January 2016 until August 2016 [8], and the evaluation of the first year of full implementation of the e-platform for the MLP was conducted from September 2016 to August 2017. The evaluation included both a qualitative [8,9] and quantitative methodology. The objective of the quantitative evaluation was to empirically explore the acceptance based on the technology acceptance model (TAM; TAM2 as a baseline model [11-13]) and information system (IS) success based on the IS success model (ISSM) [14] of the e-platform. The TAM assumes that the users' acceptance of technology depends on 2 variables: (1) perceived usefulness and (2) perceived ease of use [15]. In the original TAM, these 2 variables served as a proxy to determine the users' attitude toward using the technology [16]. In 2000, the extended TAM was proposed (TAM2) [17]to include social and cognitive influences to better reflect the complexity of the users' technology acceptance decisions and the increased complexity of more developed ISs [15]. The ISSM is a measure for IS that is organized in 3 levels: (1) the first level includes ISSM quality constructs that impact the user's satisfaction and intention to use the IS, (2) the second level examines the systems' performance, and (3) the third level looks at net benefits of an IS.

With questionnaires based on the TAM2 and ISSM, we wanted to measure technology acceptance and IS success, respectively, answering the following specific research questions:

1. Is the technology of e-learning for medical education accepted, that is, do students and medical lecturers agree with the technological environment of e-learning as a mode of learning and teaching for the CCHS?
2. To which extent is medical e-learning used for learning and teaching by MLP students and medical lecturers and how is medical e-learning used Web-based and offline (context, frequency, materials and IS success respectively)?

Methods

Study Participants

During the study period from September 2016 to August 2017, a total of 83 students (3rd study year [SY] n=23, 4th SY n=32, and bridging SY n=28) were registered in the MLP at CCHS and together with the medical lecturers (n=36) constituted the total study population. All enrolled MLP students were eligible for the study, so no sample size specification was applicable. At the time of the study, the MLP student population included 3rd and 4th SY students and bridging SY students (those who had already acquired an MLP diploma in the past and continued

1 more year in the MLP to receive a Bachelor of Science degree).

Study Design and Procedures

The prospective, nonrandomized intervention study was conducted during the MLP SY 2016/17 that started in September 2016 and ended in July 2017. The study took place in the capital—Lusaka—and 10 rural hospitals were used as practical training sites for the MLP. The CCHS main campus, a medical multiprofessional training institution, was located about 10 km east of Lusaka's city center and was one of the largest training institutions for Zambian health workers.

Learning Materials

Learning materials for the e-platform mainly comprised readily available materials provided by medical lecturers according to the MLP curriculum and materials customized for the e-platform. Access to e-platform materials was available offline via the tablets and Web-based, via the Moodle learning management system [10]. E-platform contents included static and interactive elements, such as lecture presentations (mostly Microsoft PowerPoint), medical books, interactive virtual patient cases, medical images, and short videos on medical procedures (see Table 1). Health care practice was supported with materials on standard treatment guidelines and algorithms for diagnosis and treatment.

Data Collection

Fieldwork for this study was conducted in collaboration with CCHS. In September 2016, study baseline data were collected before the intervention and included student demographic data and technology experience. Students' name, date, place, date of birth, marital status, SY, prior studies, year of graduation, and medical experience in years were collected (see Table 2) together with a 10-item paper-based questionnaire about technology experience (see Multimedia Appendix 1). A questionnaire asked students about their usage and exposure to computers and mobile computing devices, level of comfort of internet navigation, and prior experience with and self-perceived usefulness of e-learning.

After completion of the baseline data, the MLP students were given tablets (7-inch and Android-based) preloaded with learning and clinical decision-making support materials and access to the Web-based e-platform. MLP lecturers were provided access to the Web-based MLP e-platform.

In August 2017, students' data were collected on technology acceptance and IS success with paper-based questionnaires (5-item Likert scale: strongly disagree, disagree, neutral, agree, and strongly agree). The students' technology acceptance questionnaire comprised 25 questions based on the TAM2 baseline model [11-13] with 6 constructs (see Table 3): (1) perceived ease of use, (2) perceived usefulness, (3) attitude, (4) behavioral intention, (5) self-efficacy, and (6) subjective norm; and 55 questions comprised DeLone and McLean's IS success model [14] on the basis of 6 constructs (see Table 4): (1) information quality, (2) service quality, (3) system quality, (4) user satisfaction, (5) system use (the intention to use the system), and (6) net benefits. MLP students completed the questionnaires in small groups of approximately 30 students. The TAM-based questionnaire for medical lecturers included 45 questions. Medical lecturers were invited to complete the questionnaire via a cloud-based survey service within a 3-month time frame (July to October 2017), as they were dispersed throughout Zambia.

This study report conforms with the Consolidated Standards of Reporting Trials-Electronic Health [18] checklist (see Multimedia Appendix 2).

Table 1. Available Zambian medical licentiate program learning materials on the electronic platform by medical specialty and content type.

Content type	Medical specialty				Total
	Internal medicine	Surgery	Obstetrics and gynecology	Pediatrics	
Lecture notes	38	64	43	31	176
Medical books	6	1	1	7	15
Exam preparation	1	0	0	1	2
Treatment guidelines	69	0	1	32	102
Videos	2	44	42	0	88
Pictures	0	0	0	39	39
Virtual patients	4	1	0	2	7

Table 2. Demographic data (summary statistics) of the study sample of medical licentiate practitioner (MLP) students and lecturers.

Variables	n (%)	Minimum	Maximum	Median	Mean (SD)
MLP students (N=74)					
Gender		__[a]	—	—	—
Female	16 (22)				
Male	58 (78)				
Age (years)		**23**	**54**	**36**	**38 (7)**
≤30	10 (14)				
>30-≤40	42 (57)				
>40	22 (30)				
Medical experience (years)		**1**	**28**	**9**	**10 (6)**
≤5	21 (28)				
>5-≤15	35 (47)				
>15	18 (24)				
Technology experience score between 0 (minimum) and 1 (maximum)		**0.06**	**0.88**	**0.69**	**0.67 (0.16)**
Low	27 (36)				
Moderate	36 (49)				
High	11 (15)				
Study year		—	—	—	—
Third	20 (27)				
Fourth	29 (39)				
Bridging	25 (34)				
MLP medical lecturers (N=14)					
Gender		—	—	—	—
Female	4 (29)				
Male	10 (71)				
Age (years)		**38**	**54**	—	**45 (6)**
≤30	0 (0)				
>30-≤40	3 (21)				
>40	11 (79)				
Medical experience (years)		**9**	**27**	**14.5**	**17 (6)**
≤5	0 (0)				
>5-≤15	8 (57)				
>16	6 (43)				
Teaching experience (years)		**2**	**2**	—	**10 (6)**
≤5	4 (29)				
>5-≤15	8 (57)				
>15	2 (14)				

[a]Not applicable.

Table 3. Questionnaire results of technology assessment model questionnaires of medical licentiate practitioner students.

Questionnaire items of technology acceptance component: MLP[a] students	Mean (SD)	r [b]
Perceived ease of use	1.82 (0.61)	0.61
I find the MLP e-learning platform easy to use.	1.89 (0.57)	0.45
Learning how to use the MLP e-learning platform is easy for me.	1.96 (0.60)	0.57
It is easy to become skillful at using the MLP e-learning platform.	1.90 (0.64)	0.63
Learning to operate my tablet is easy for me.	1.80 (0.71)	0.61
I am clear on how to use the tablet.	1.77 (0.64)	0.65
It is easy for me to become skillful at using my tablet.	1.68 (0.53)	0.69
I find my tablet easy to use.	1.73 (0.51)	0.65
Perceived usefulness	1.85 (0.61)	0.60
E-learning improves my learning performance.	1.93 (0.49)	0.49
E-learning makes it easier to study course content.	1.96 (0.62)	0.49
Using a tablet computer is compatible with all aspects of my studies.	2.04 (0.78)	0.64
I think that using a tablet fits well with the way I like to learn.	1.90 (0.61)	0.73
Using a tablet fits into my learning style.	1.83 (0.56)	0.66
In my job, using a tablet is important.	1.61 (0.52)	0.64
In my job, using e-learning is important.	1.65 (0.54)	0.50
Attitude	1.72 (0.59)	0.70
Studying through e-learning is a good idea.	1.79 (0.59)	0.73
Studying through e-learning is a wise idea.	1.76 (0.60)	0.75
I am positive towards e-learning.	1.74 (0.61)	0.72
I am positive towards using a tablet for medical learning.	1.60 (0.57)	0.59
Behavioral intention	1.69 (0.58)	0.67
I intend to be a heavy user of the MLP e-learning platform.	1.80 (0.63)	0.55
For my future job, it is necessary to know how to use a tablet.	1.67 (0.56)	0.69
For my future job, it is necessary to know how to use a computer.	1.59 (0.55)	0.77
Self-efficacy	1.86 (0.69)	0.65
I feel confident finding information on the MLP e-learning platform.	1.87 (0.63)	0.72
I have the necessary skills for using the MLP e-learning platform.	1.94 (0.67)	0.65
I feel confident using the MLP e-learning (Moodle app) with the tablet.	1.77 (0.76)	0.59
Subjective norm	1.92 (0.61)	0.78
What e-learning stands for is important for me as an MLP student.	1.90 (0.61)	0.75
It is necessary to take e-learning courses to train as an MLP.	1.93 (0.61)	0.80
Information system success (for details see information system success model results)	2.13 (0.76)	0.29

[a]MLP: medical licentiate practitioner.

[b]r: corrected item-total correlations.

Table 4. Questionnaire results of information system success of medical licentiate practitioner students.

Questionnaire information system success component: MLP[a] students	Mean (SD)	r [b]
Information quality	1.99 (0.80)	0.57
The MLP e-learning provides up-to-date learning materials.	2.06 (0.86)	0.58
Learning materials available on the MLP e-learning platform are clear.	1.88 (0.72)	0.59
Materials available on the MLP e-learning platform are useful to me.	1.58 (0.56)	0.49
Materials available on the MLP e-learning platform help me to understand medical topics better.	1.89 (0.69)	0.57
The quality of materials available on the MLP e-learning platform is high.	2.34 (0.85)	0.66
The MLP e-learning provides information relevant to MLP medical practice.	1.63 (0.55)	0.56
Materials on the MLP e-learning platform increase my quality of clinical care.	1.74 (0.51)	0.57
The MLP e-learning platform provides sufficient learning materials.	2.51 (0.94)	0.65
Through the MLP e-learning platform. I get access to learning materials I need in time.	2.29 (0.88)	0.48
Service quality	2.29 (0.86)	0.52
The MLP e-learning platform provides proper level of assistance and explanation.	2.30 (0.74)	0.37
The MLP e-learning platform was available when I wanted to access it.	2.65 (1.00)	0.33
The local IT[c] has adequate knowledge to help me if I experience any problems with the MLP e-learning platform.	2.28 (0.88)	0.66
The local IT has adequate knowledge to help me if I experience any problems with the tablet.	2.07 (0.74)	0.62
The IT provides satisfactory support to users of the MLP e-learning platform.	2.11 (0.87)	0.63
The IT attends to my problems.	2.15 (0.84)	0.61
The MLP e-learning platform provides dependable services.	2.24 (0.73)	0.52
The MLP e-learning platform provides rapid services.	2.54 (0.92)	0.38
System quality	2.10 (0.70)	0.61
The MLP e-learning platform is easy to use.	2.02 (0.77)	0.66
The MLP e-learning platform is user-friendly.	1.94 (0.72)	0.69
The MLP e-learning platform is easy to learn.	1.89 (0.61)	0.58
Most MLP students find the MLP e-learning platform easy to use.	2.15 (0.66)	0.66
Most MLP students find the MLP e-learning platform user-friendly.	2.23 (0.74)	0.65
Most MLP students find the MLP e-learning platform easy to learn.	2.14 (0.65)	0.61
The user interface of the MLP e-learning is attractive.	2.17 (0.69)	0.49
The MLP e-learning platform has attractive features to appeal to users.	2.23 (0.72)	0.56
User satisfaction	2.11 (0.72)	0.62
I think that most MLP students bring a positive attitude towards the MLP e-learning platform.	2.11 (0.61)	0.51
I think that most MLP students have a high perceived utility about the MLP e-learning platform.	2.22 (0.63)	0.64
I am satisfied with efficiency of the MLP e-learning platform.	2.30 (0.88)	0.70
I will continue to use the MLP e-learning platform.	1.79 (0.67)	0.59
Overall, I am very satisfied with the MLP e-learning platform.	2.14 (0.67)	0.66
System use	2.47 (0.84)	0.69
My frequency of using the MLP e-learning platform is high.	2.33 (0.75)	0.67
I use the MLP e-learning platform daily several times.	2.61 (0.90)	0.76
I depend upon the MLP e-learning platform.	2.53 (0.82)	0.72
I use the Moodle app daily several times.	2.40 (0.67)	0.60
Net benefits	2.00 (0.61)	0.60
The MLP e-learning platform helps me prepare better for the MLP exam.	2.15 (0.76)	0.64

Questionnaire information system success component: MLP[a] students	Mean (SD)	r [b]
The MLP e-learning platform helps me think through medical problems.	2.11 (0.59)	0.54
Using the MLP e-learning platform has helped me to accomplish my learning tasks more efficiently.	2.25 (0.62)	0.66
Using the MLP e-learning platform has made my learning activities become much easier than without.	2.14 (0.66)	0.63
My learning performance [is] enhanced since using the e-learning platform.	2.04 (0.62)	0.63
I find the MLP e-learning platform useful for my studies.	1.92 (0.55)	0.59
The MLP e-learning platform saves me money.	1.96 (0.66)	0.62
The e-learning materials improve my clinical performance.	1.93 (0.54)	0.60
I feel that the e-learning platform has a direct positive impact on being an MLP practitioner.	1.83 (0.53)	0.56
The MLP e-learning platform helps me to treat patients better.	1.82 (0.48)	0.53
The MLP e-learning platform saves me time.	1.90 (0.56)	0.62

[a]MLP: medical licentiate practitioner.

[b]r: corrected item-total correlations.

[c]IT: information technology.

Objectives

The primary objective of the study was to evaluate the technology acceptance and IS success of the e-platform.

Statistical Methods

All data were cleaned in a systematic screening for completeness, plausibility, and consistency. Potential inconsistencies were resolved by checking against original data forms. Results were transferred into a spreadsheet and data were then analyzed with RStudio Desktop (open-source license version 1.1.463) [19].

For questionnaire items, we calculated the mean, standard deviation, and correlation coefficient r to see how well aligned questions were to the respective categories. Likert scale data were interpreted as interval scales [20], and data from various study participants were regarded as independent as we assumed that the behavior of one participant did not influence the behavior of another. Correlations of variables were explored with Kendall tau correlation coefficient as an estimator of correlation in the population.

Ethical Considerations

We communicated the purpose of this research to study participants and explained the study design and participation requirements. Before taking part in this study, all study participants agreed to and signed an informed consent explaining the study scope and purpose, and their right to withdraw at any point. Voluntary study participation was emphasized to be voluntary and assessments were solely part of this study. Furthermore, it was emphasized that the usage of the e-platform is tracked. The Biomedical Research Ethics Committee of the University of Zambia and the ethical committee of the University Hospital Heidelberg, Germany approved the study protocol.

Results

Demographics

Overall, the student population was predominantly male (58/74; 78%) with female students in the minority (16/74; 22%). The predominant age group was between 30 to 40 years (42/74; 57%; see Table 2). Most students had had more than five years of medical experience with an average of 10 years of working experience. Medical lecturers in the MLP were also predominantly male (10/14; 71%) with an average work experience of approximately 17 years and an average of 10 years of teaching experience.

Female students' experience with technology was generally low, whereas males reported a predominantly midrange of technology experience (see Multimedia Appendix 3). Demographic details for age groups and medical and technology experience by gender are presented in Table 2.

Owing to a high number of missing data values, data from 9 MLP students were excluded from further analysis.

Technology Acceptance

With regard to technology acceptance, the highest rated categories by MLP students were *Behavioral Intention* (1.69; SD 0.58; r=0.67) and *Attitude* (1.72 SD=0.59; r=0.70), whereas *Subjective Norm* (1.92; SD=0.61; r=0.78) and *Self-Efficacy* (1.86; SD=0.69; r=0.65) were the categories with the lowest acceptance among students, followed by the categories of perceived usefulness (=1.85; SD=0.61; r=0.60) and perceived ease of use (=1.82; SD=0.61; r=0.61; see Table 3). The individual items with the highest agreement among MLP students were as follows: "For my future job, it is necessary to know how to use a computer" (item 21; 1.59; SD=0.55; r=0.77), "I am positive towards using a tablet for medical learning" (item 18; 1.60; SD=0.57; r=0.59), and "In

my job, using a tablet is important" (item 13; 1.61; SD=0.52; *r*=0.64). The least agreement MLP students had were on the following items: "Using a tablet computer is compatible with all aspects of my studies" (item 10; 2.04; SD=0.78; *r*=0.64), "E-learning makes it easier to study course content" (item 9; 1.96; SD=0.62; *r*=0.49), and "Learning how to use the MLP e-learning platform is easy for me" (item 2; 1.96; SD=0.60; *r*=0.57).

Information System Success

For the IS success questionnaire, the category of Information Quality received the overall best ratings from MLP students (1.99; SD=0.80; *r*=0.57) followed by Net Benefits (2.00; SD=0.61; *r*=0.60) that contrasted with System Use with the lowest rating (2.47; SD=0.84; *r*=0.69) and Service Quality (2.29; SD=0.86; *r*=0.52) with the second lowest rating (see Tables 3 and 4). The specific items with the highest agreement among MLP students were all in the category of information quality, "Materials available on the MLP e-learning platform are useful to me" (item 3; 1.58; SD=0.56; *r*=0.49), "The MLP e-learning provides information relevant to MLP medical practice" (item 6; 1.63; SD=0.55; *r*=0.56), and "Materials on the MLP e-learning platform increase my quality of clinical care" (item 7; 1.74; SD=0.51; *r*=0.57). Items rated the lowest by MLP students were in the lowest-rated categories of system use and service quality, "I use the MLP e-learning platform daily several times" (item 32; 2.61; SD=0.90; *r*=0.76), "I depend upon the MLP e-learning platform" (item 17; 2.54; SD=0.92; *r*=0.38), and "The MLP e-learning platform provides rapid services" (item 33; 2.53; SD=0.82; *r*=0.72).

The Kendall tau variable correlation coefficients are shown in Tables 5 and 6 for the IS success model and TAM questionnaires, respectively. The bivariate relationship indicated that all of the variables (questionnaire items) were significantly correlated (correlations<0.05; see Tables 5 and 6).

For medical lecturers, the categories *Perceived enjoyment* (1.34; SD=0.48; *r*=0.73) and *Perceived ease of use* (1.86; SD=1.05; *r*=0.48) were rated the highest, whereas *Behavioral intention* (2.14; SD=0.83; *r*=0.54) and *Self-efficacy* (2.11; SD=1.19; *r*=0.72) were perceived as the categories with the least agreement (see Table 7). The Kendall tau variable correlation coefficients for the TAM questionnaire are shown in Table 8. Furthermore, the questions asked in addition to the TAM disclosed that 2 medical teachers had never used the e-platform (never n=2; often n=3; once n=1; several times n=1; frequently n=1) and 5 did not know how to access it. Overall, 50% (7/14) of medical lecturers stated that they had contributed to the e-platform with content or other e-learning–based activities, such as virtual patients (n=7). Correlation coefficients (Kendall ; nonparametric correlation) of the medical lecturers' technology acceptance items indicated a correlation for most variables (see Table 8). High correlation significance (*P*<.01) was shown between the items perceived usefulness and attitude, perceived ease of use and self-efficacy, and perceived ease of use and perceived usefulness.

Table 5. Correlations (Kendall τ) of information system success model of the Zambian medical licentiate practitioner student questionnaire.

Categories of information system success model	Service quality (ServQ)	System quality (SQ)	User satisfaction (US)	System use (SU)	Net benefits (NB)
Information quality (InfQ)	0.23[a]	0.36[b]	0.41[b]	0.30[b]	0.40[b]
ServQ	—[c]	0.35[b]	0.25[a]	0.26[a]	0.22[a]
SQ	—	—	0.49[b]	0.31[b]	0.37[b]
US	—	—	—	0.40[b]	0.61[b]
SU	—	—	—	—	0.46[b]
NB	—	—	—	—	—

[a]*P*<.05.

[b]*P*<.01.

[c]Not applicable.

Table 6. Correlations (Kendall τ) of adapted technology acceptance model of the Zambian medical licentiate practitioner student questionnaire.

Categories of technology acceptance model	Perceived usefulness (PU)	Attitude (AT)	Behavioral intention (BI)	Self-efficacy (SE)	Subjective norm (SN)	Information system success (IS)
Perceived ease of use	0.48[a]	0.46[a]	0.48[a]	0.31[b]	0.31[b]	0.36[a]
PU	—[c]	0.53[a]	0.41[a]	0.52[a]	0.47[a]	0.44[a]
AT	—	—	0.59[a]	0.44[a]	0.39[a]	0.35[a]
BI	—	—	—	0.47[a]	0.37[a]	0.22[b]
SE	—	—	—	—	0.46[a]	0.36[a]
SN	—	—	—	—	—	0.38[a]
IS	—	—	—	—	—	—

[a]*P*<.01.

[b]*P*<.05.

[c]Not applicable.

Table 7. Questionnaire results of technology acceptance model questionnaires of medical licentiate practitioner lecturers.

Questionnaire items of technology acceptance component: MLP[a] medical lecturers	Mean (SD)	r [b]
Attitude	1.90 (0.80)	0.39
I intend to use the MLP e-learning platform to assist my medical teaching.	1.50 (0.67)	0.37
I am positive towards the MLP e-learning platform.	1.92 (0.79)	0.40
I believe that working with tablets is for young people only.	1.25 (0.45)	0.42
I believe that working with computers is for young people only.	1.25 (0.45)	0.42
Most MLP medical lecturers bring a positive attitude towards e-learning.	2.58 (0.67)	0.50
I want to dedicate more effort to support the MLP e-learning platform.	2.08 (1.08)	0.37
The MLP e-learning platform can be successfully established for the MLP program.	2.08 (0.67)	0.36
I think the MLP e-learning platform can be established as a permanent part of the MLP program.	2.08 (0.67)	0.36
E-learning is a good tool for the MLP program at CCHS.	1.83 (0.72)	0.36
I think the MLP e-learning platform is important for MLP students.	2.08 (0.79)	0.38
Behavioral intention	2.14 (0.83)	0.54
I want to upload content myself on the MLP e-learning platform.	2.36 (1.03)	0.42
My input is vital for the success of the MLP e-learning platform.	2.18 (0.75)	0.41
I actively want to contribute to the MLP e-learning platform.	1.91 (0.54)	0.41
I think other MLP medical lecturers want to regularly contribute to the MLP e-learning platform.	2.45 (0.93)	0.39
I want to regularly contribute content to the MLP e-learning platform.	2.09 (1.14)	0.50
I am willing to take ownership for the MLP e-learning platform.	2.09 (0.83)	0.37
The MLP e-learning platform is only temporary.	2.27 (0.79)	0.37
The MLP e-learning platform can be sustained long-term at CCHS.	2.27 (0.79)	0.37
I think other MLP medical lecturers want the MLP e-learning platform to be successful.	1.91 (0.83)	0.36
I want the MLP e-learning platform to be successful.	1.82 (0.60)	0.35
Perceived ease of use	1.86 (1.05)	0.48
Working with computers is easy for me.	1.36 (0.50)	0.07
I am using computers as a tool for teaching.	1.36 (0.50)	0.04
Learning how to use the MLP e-learning platform is easy for me.	2.36 (0.92)	0.16
I need training to actively contribute to the MLP e-learning platform.	2.36 (1.50)	0.22
Perceived usefulness	2.00 (0.84)	0.45
The use of The MLP e-learning platform is better as compared to textbook learning.	2.45 (0.82)	0.45
Overall, the use of The MLP e-learning platform is more time-demanding than traditional teaching methods.	2.00 (1.26)	0.49
The MLP e-learning platform provides quality medical materials for MLP students.	2.00 (0.77)	0.44
The MLP e-learning platform is a useful tool for the MLP program.	1.82 (0.75)	0.43
I think e-learning platform improves my productivity as a medical lecturer.	2.09 (1.04)	0.45
I think e-learning improves my effectiveness as a medical lecturer.	1.82 (0.98)	0.43
I think e-learning improves my teaching performance.	1.82 (0.98)	0.43
I believe using e-learning is helpful for my teaching.	1.64 (0.81)	0.44
E-learning makes it easier to teach medical courses and their content.	1.91 (0.70)	0.43
I believe working with computers makes a person more productive at their job.	1.73 (1.01)	0.46
I think the MLP e-learning helps students to be better MLPs.	2.27 (0.65)	0.45
I think that the MLP e-learning can improve quality of learning for MLP students.	1.73 (0.65)	0.44
I think the MLP e-learning platform needs too much effort to use it for teaching.	2.27 (0.90)	0.51

Questionnaire items of technology acceptance component: MLP[a] medical lecturers	Mean (SD)	*r* [b]
I think the MLP e-learning platform improves the effectiveness of medical teaching.	1.91 (0.70)	0.44
I think the MLP e-learning platform can save me effort.	1.82 (0.75)	0.47
Overall, I think the MLP e-learning platform can save me time.	1.91 (0.70)	0.45
The MLP e-learning platform improves medical performance of MLP students as MLP practitioners.	2.45 (0.82)	0.46
The MLP e-learning platform improves the clinical performance of MLP students.	2.18 (0.75)	0.44
Tablets are useful as learning devices for MLP students.	2.00 (0.77)	0.44
The MLP e-learning platform improves learning outcomes of MLP students.	2.18 (0.87)	0.44

[a]MLP: medical licentiate practitioner.

[b]*r*: item-total correlation.

[c]Reverse scoring of Likert items applied.

Table 8. Correlation (Kendall τ) of adapted technology acceptance model of Zambian medical licentiate practitioner lecturer questionnaire.

Categories of technology acceptance model	Perceived usefulness (PU)	Attitude (AT)	Behavioral intention (BI)	Self-efficacy (SE)
Perceived ease of use	0.40[a]	0.53[b]	-0.02	0.62[b]
PU	—[c]	0.75[d]	0.44[a]	0.60[b]
AT	—[c]	—[c]	0.32	0.53[b]
BI	—[c]	—[c]	—[c]	0.11
SE	—[c]	—[c]	—[c]	—[c]

[a]*P*<.10.

[b]*P*<.05.

[c]Not applicable.

[d]*P*<.001.

Discussion

Principal Findings

Overall, the questionnaires based on the TAM and IS success model indicated a positive reception of the e-platform and its usage with the offline-based tablets by MLP students and medical lecturers and identified strengths and shortcomings in the tablets and MLP e-platform. The questionnaires' results elicited an overall acceptant attitude toward the e-platform and agreement with the technological environment of e-learning as a mode of learning and teaching by both, MLP students and medical lecturers.

The student study participants were quite heterogeneous in age, medical experience, and technological experience. The youngest student was aged 23 years and the oldest was 54 years. Some students had only 1 year of medical experience whereas others had accumulated over 28 years of medical experience. The same was applied to technological experience, as some students' exposure to the MLP e-platform was the first time they had ever had a chance to experience continuous technology use for learning and accessing knowledge. Others, *digital natives* who had grown up with technologies, came into the MLP program equipped with 2 mobile devices and a laptop as an inherent part of their daily lives. Males seemed more confident in rating their self-perceived technology experience than females, who seemed more conservative in their self-assessment. Potentially, this finding may be the result of the prevailing gender-based stereotypes holding females back, as technology usage may not fit the stereotype of expectations for African females. In general, female students may be faced with "a subconscious or implicit bias against women scientists in higher education settings" as "stereotypes emerge early and continue to be salient throughout the lifespan" [21]. The implicit bias against women, which is unfortunately still prevalent, should be taken up and actively approached in the MLP e-platform's implementation strategy. One study found, "for instance, college students were more likely to rate the same conference abstracts as lower in scientific quality if the author's name was female instead of male, particularly if these topics had traditional masculine themes" [21]. Another study described a "moderating effect of gender difference in the adoption of multimedia technology for learning," whereby males, "believe the technology can improve their performance when they find a fit between task and technology, and this belief would lead to their adoption of the technology" [22]. The study recommended (only for males) "that educators can better promote male students' adoption of multimedia technology by demonstrating how appropriate the multimedia technology given is for their learning" [22].

Age too can be a discriminative factor in technology acceptance and should be taken into consideration particularly for heterogeneous cohorts like the MLP students. "More attention should be paid to their support [elder MLP students] in the form of training, as well as meeting their unique personal needs"

[23]. Overall, elder MLP students were more satisfied with the MLP e-platform than younger MLP students. The higher satisfaction may have been rooted in older students' limited prior exposure to technology and consequently lower expectation levels. Thus, elder MLP students may have perceived the MLP e-platform as more satisfying than their younger colleagues.

In particular, MLP e-platform information quality and net benefits were perceived to be the most appealing characteristics based on the IS success model. A closer look at the individual questions of those 2 categories, however, shows that the ratings were more differentiated since students rated available materials as clear and useful, but questions about materials' quality and sufficiency received the lowest overall scores (information quality). The quality and quantity of contents was perceived as a hindrance which may have been rooted in the short implementation period of the MLP e-platform in general which resulted in limited and rather static learning content that did not entirely reflect the MLP curriculum. Contents were initially composed only of lecture notes and adapted materials from a Malawian e-learning platform [24,25].

To this end, adopting a social media style for content creation and curation to allow for student-made, peer-reviewed content based on fellow student ratings, mixed with methods such as microlearning or learning nuggets, could potentially leverage the practicality of mobile devices (eg, tablets), and could be a significant next step for higher acceptance learning [26]. Curating available Web-based content may be of benefit to keep contents up to date while at the same time decreasing efforts to develop own learning content which constitutes a current major bottleneck—resource-intense and manual content creation. The MLP e-platform may then act also as a central hub for grouping and sharing reliable and appropriate information from the internet [27].

Students underlined with their high acceptance of the statements "For my future job, it is necessary to know how to use a computer," "I am positive towards using a tablet for medical learning," and "In my job, using a tablet is important" that they are willing and know the importance to make use of technologies, such as tablets, for their clinical work. Consequently, the categories of perceived usefulness denoting the relevance of the MLP e-platform toward job performance and perceived ease of use also had high agreement among MLP students. Similarly, for e-platform net benefits, the students found it useful in general, especially using the tablet in a clinical context for treating patients, but the e-platform did not support students well for their exams or help them achieve their learning tasks more efficiently. The subjective norm, *a person's perception that most people who are important to him think he should or should not perform the behavior in question* [28], was not rated as a strong factor compared with other questionnaire categories nor was e-platform self-efficacy, *people's judgments of their capabilities to perform a given task* [29]. The students' low ratings for the subjective norm can be attributed to the e-platform still being a relatively new mode for the MLP. The MLP students did not find the tablet compatible with all aspects of their studies, nor did they perceive that the e-platform made studying easier. This shows a need for adaptation of the e-platform to better suit its users and enhance the overall ease of use. During the SY, tablet failure impacted the reliable usage of the offline e-platform (service quality). Most tablets failed either because of poor operating system upgrades or hardware fragility (for a more detailed description see the study by Barteit et al [9]). Furthermore, for some tablets, the battery life span reduced rapidly limiting its usage. Service quality and the quality of the materials, in turn, may have had an impact on the students' low ratings for self-efficacy, which has been shown to have *a direct and powerful effect on actual use over and above user intention* [29].

Medical lecturers showed a general acceptance of technology for learning and teaching within the MLP program as the highest ratings for the questionnaires were given for perceived ease of use and perceived usefulness of the e-platform. However, most medical lecturers rated their colleagues' attitude toward the MLP e-platform as not very positive, which indicates that many did not count on digitized methods to be a part of their teaching and potentially do not see these methods as valid instruments. Their perception toward the e-platform to improve MLP student medical performance received low agreement scores, which further underlines skepticism toward e-learning as a teaching method. In addition, medical lecturers conveyed low confidence toward using the e-platform for medical teaching in combination with a potentially conservative preference for textbook-based learning and lecture-style teaching as a core teaching method. The item, *The use of the ML e-learning platform is better as compared to textbook learning*, received one of the lowest ratings of all questions, and the categories of behavioral intention and self-efficacy of the e-platform both received very low scores indicating low agreement among medical lecturers. These low ratings further manifest a certain reluctance to employ the e-platform for teaching. Potentially, a bias was rooted in social desirability which influenced MLP lecturers' answers to the questionnaire [30] as they knew that they were expected to show a certain openness to technologies. Thus, the results of this evaluation suggest a low medical lecturer uptake of the e-platform, which potentially resulted in limited learning materials and restricted integration of the e-platform in day-to-day teaching and training [9,31]. One MLP medical lecturer suggested increasing the allowance for *site consultants so as to motivate them* to make use of the e-platform. In fact, an increase in lecturer payment in Mexico showed no benefit, as it did not translate into an impact on student learning nor did it hold as a strategy for better teacher engagement in the digital learning and teaching process [32]. A more promising approach seems to be to consider the transformation of the *teaching-learning process* that entails a global approval and acceptance of teachers and the administration of digital technologies. Technology cannot fix outdated processes and substitute for well-qualified, motivated teachers. Thus, it may be of benefit to further educate and support teachers to use digital technologies for teaching and training [31], including content creation and e-learning didactics, such as flipped classrooms, adaptive learning [33], video-making with low-entry technologies such as smartphones, short mobile learning sessions (microlearning for on-the-job learning), and fostering technology-based social learning activities such as discussion and learning groups. An increased institutional presence may also improve lecturer approval and acceptance, such as an

e-learning lab, which may take the physical shape of a dedicated part of the library with a few computers, where medical lecturers can seek technical support for making use of the e-platform or to prepare e-platform contents. Attitude has been identified as an important factor contributing to technology acceptance and may well change over time when more technological experience is gained [34]. Training teachers in digital literacy is not only of benefit for catering for the *teaching-learning process* but also may provide the cobenefit of preparing them to handle and embrace technologies in the medical field, which are pervasive and rapidly increasing as the digitization of medicine progresses [35,36].

Overall, medical lecturers did show a readiness and acceptance toward technology, since the majority negated that technologies were only for young people and agreed that they intended to use the e-platform to assist their medical teaching and that the e-platform could be helpful for their teaching.

Limitations

The limitations of the study are as follows:

- Ensuring questions are clear and not misleading: (1) Getting respondents to answer questions thoughtfully and honestly and (2) obtaining a sufficient number of completed questionnaires to enable meaningful analyses [37].
- No comparison group: All study participants were subject to the same intervention, so it was difficult to attribute the change in outcome to the introduction of the program with any certainty.

Conclusions

MLP students and medical lecturers accepted the e-platform as a method for medical teaching and learning and the evaluation of the MLP e-platform with its offline tablet-based component, which proved to be a feasible approach for teaching and learning within the low-resource environment at the CCHS in Zambia. This overall positive acceptance toward the e-platform constitutes a fertile base to scale up and implement the e-platform as a serious learning and teaching methodology within the blended learning strategy for the MLP program and beyond. Main shortcomings comprise the low uptake of the e-platform in everyday teaching and the scope and level of interactivity of the e-platform contents. The e-platform has the potential to be a cornerstone in the expansion of scaling up the training of urgently needed medical licentiates in Zambia, but it requires a profound transformation of the teaching and learning process that at its core is manifested in the curriculum and ongoing technology training for medical lecturers. From a technological standpoint, the tablet-based e-platform is a flexible tool that allows for electronic assessment, skill-based learning sessions, a digital logbook for the MLP, and an enhanced clinical decision-support system that could be incorporated in the continuous medical training for MLP graduates in rural health facilities. With increasingly affordable and viable technologies being available, the e-platform may potentially become a game changer in future medical education, especially in a low-resource context such as Zambia.

Acknowledgments

The authors wish to thank Else Kröner-Fresenius-Stiftung [2016_HA67] for supporting this study. The authors would like to acknowledge SolidarMed that has supported the MLP program since 2009 for their openness to collaborate, which provided the base for this partnership project. The authors also thank Julia M Challinor for her editorial support.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Questionnaire for medical licentiate practitioner students comprising demographic data and prior technological experience.
[PDF File (Adobe PDF File), 1278 KB - jmir_v21i10e14748_app1.pdf]

Multimedia Appendix 2
CONSORT eHealth Checklist.
[PDF File (Adobe PDF File), 359 KB - jmir_v21i10e14748_app2.pdf]

Multimedia Appendix 3
Level of technology experience amongst medical licentiate practitioner (MLP) student study population by gender.
[PNG File , 486 KB - jmir_v21i10e14748_app3.png]

References

1. Ministry of Health. Zambia National Health Strategic Plan 2017-2021 Internet. In: National Development Plan. Lusaka, Zambia: Ministry of Health; 2017.
2. Ministry of Health. Human Resources for Health Planning and Development Strategy Framework. In: National Development Plan. Lusaka, Zambia: Ministry of Health; 2017.

3. Barteit S. Scribd. Ministry of Health. National Human Resources for Health Strategic Plan 2018-2024. Lusaka, Zambia: Ministry of Health; 2018 URL: http://www.moh.gov.zm/?page_id=5620 [accessed 2019-09-24]
4. World Health Organization. Health workforce requirements for universal health coverage and the Sustainable Development Goals. Geneva, Switzerland: World Health Organization; 2016.
5. World Bank Data. The World Bank In Zambia. URL: https://www.worldbank.org/en/country/zambia [accessed 2018-11-25]
6. Graf U, Bowa A, Thieme K, Häggblom A. Strengthening Zambia's Health Workforce: The Medical Licentiate Training Programme. In: Proceedings of the Geneva Health Forum. Presented at: Geneva Health Forum; 2014; Geneva, Switzerland URL: https://www.researchgate.net/publication/262418301_Strengthening_Zambia's_health_workforce_the_Medical_Licentiate_training_programme
7. Eysenbach G. What is e-health? J Med Internet Res 2001;3(2):E20 [FREE Full text] [doi: 10.2196/jmir.3.2.e20] [Medline: 11720962]
8. Barteit S, Jahn A, Bowa A, Lüders S, Malunga G, Marimo C, et al. How self-directed e-learning contributes to training for medical licentiate practitioners in Zambia: evaluation of the pilot phase of a mixed-methods study. JMIR Med Educ 2018 Nov 27;4(2):e10222 [FREE Full text] [doi: 10.2196/10222] [Medline: 30482744]
9. Barteit S, Neuhann F, Bärnighausen T, Bowa A, Lüders S, Malunga G, et al. Perspectives of nonphysician clinical students and medical lecturers on tablet-based health care practice support for medical education in Zambia, Africa: qualitative study. JMIR Mhealth Uhealth 2019 Jan 15;7(1):e12637 [FREE Full text] [doi: 10.2196/12637] [Medline: 30664475]
10. Moodle. 2019. URL: https://www.moodle.org [accessed 2019-09-20]
11. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q JSTOR 1989 Sep;13(3):319-319. [doi: 10.2307/249008]
12. Davis FD. 1985. A Technology Acceptance Model for Empirically Testing New End-User Information Systems: Theory and Results. Massachusetts Institute of Technology; 1985 URL: https://dspace.mit.edu/handle/1721.1/15192 [accessed 2019-09-24]
13. Park SY. An analysis of the technology acceptance model in understanding university students' behavioral intention to use e-learning. Educ Technol Soc JSTOR 2009;12(3):162.
14. Delone WH, McLean ER. The DeLone and McLean model of information systems success: a ten-year update. J Manag Inf Syst 2003;19(4):9-30. [doi: 10.1080/07421222.2003.11045748]
15. Lee Y, Kozar KA, Larsen KR. The technology acceptance model: past, present, and future. Commun Assoc Inf Syst 2003 May;12(5):924-8, 928.e1. [doi: 10.17705/1CAIS.01250]
16. Chuttur MY. Overview of the technology acceptance model: origins, developments and future directions. Sprouts Work Pap Inf Syst 2009;9(37):1-22 [FREE Full text]
17. Venkatesh V, Davis FD. A theoretical extension of the technology acceptance model: four longitudinal field studies. Manage Sci 2000 Feb;46(2):186-204. [doi: 10.1287/mnsc.46.2.186.11926]
18. Eysenbach G, CONSORT-EHEALTH Group. CONSORT-EHEALTH: improving and standardizing evaluation reports of Web-based and mobile health interventions. J Med Internet Res 2011 Dec 31;13(4):e126 [FREE Full text] [doi: 10.2196/jmir.1923] [Medline: 22209829]
19. RStudio. URL: http://www.rstudio.com/ [accessed 2019-01-19]
20. Sullivan GM, Artino AR. Analyzing and interpreting data from likert-type scales. J Grad Med Educ 2013 Dec;5(4):541-542 [FREE Full text] [doi: 10.4300/JGME-5-4-18] [Medline: 24454995]
21. Wang M, Degol JL. Gender gap in Science, Technology, Engineering, and Mathematics (STEM): Current knowledge, implications for practice, policy, and future directions. Educ Psychol Rev 2017 Mar;29(1):119-140 [FREE Full text] [doi: 10.1007/s10648-015-9355-x] [Medline: 28458499]
22. Park C, Kim D, Cho S, Han H. Adoption of multimedia technology for learning and gender difference. Comput Human Behav 2019 Mar;92:288-296. [doi: 10.1016/j.chb.2018.11.029]
23. Klimova B, Poulova P. Older People and Technology Acceptance. In: Proceedings for the International Conference on Human Aspects of IT for the Aged Population. 2018 Presented at: ITAP'18; July 15–20, 2018; Las Vegas, NV, USA p. 85-94. [doi: 10.1007/978-3-319-92034-4_7]
24. Barteit S, Hoepffner P, Huwendiek S, Karamagi A, Munthali C, Theurer A, et al. Self-directed e-learning at a tertiary hospital in Malawi--a qualitative evaluation and lessons learnt. GMS Z Med Ausbild 2015;32(1):Doc7 [FREE Full text] [doi: 10.3205/zma000949] [Medline: 25699110]
25. Neuhann F, Barteit S. Lessons learnt from the MAGNET Malawian-German Hospital Partnership: the German perspective on contributions to patient care and capacity development. Global Health 2017 Jul 26;13(1):50 [FREE Full text] [doi: 10.1186/s12992-017-0270-4] [Medline: 28747207]
26. Wong L, Milrad M, Specht M. Seamless Learning in the Age of Mobile Connectivity. Singapore: Springer; 2015.
27. Mihailidis P, Cohen J. Exploring Curation as a core competency in digital and media literacy education. J Interact Media Educ 2013 Feb 8;2013(1):2. [doi: 10.5334/2013-02]
28. Hill R, Fishbein M, Ajzen I. Belief, Attitude, Intention and Behavior: An Introduction to Theory and Research. Boston: Addison-Wesley; 1977.

29. Yi MY, Hwang Y. Predicting the use of web-based information systems: self-efficacy, enjoyment, learning goal orientation, and the technology acceptance model. Int J Hum Comput Stud 2003 Oct;59(4):431-449. [doi: 10.1016/s1071-5819(03)00114-9]
30. Krumpal I. Determinants of social desirability bias in sensitive surveys: a literature review. Qual Quant 2013;47(4):2025-2047. [doi: 10.1007/s11135-011-9640-9]
31. de Melo G, Machado A, Miranda A. IZA - Institute of Labor Economics. 2014. The Impact of a One Laptop per Child Program on Learning: Evidence from Uruguay URL: http://ftp.iza.org/dp8489.pdf [accessed 2019-09-24]
32. Behrman JR, Parker S, Todd P, Wolpin KI. Aligning learning incentives of students and teachers: results from a social experiment in Mexican high schools. SSRN J 2015;123(2):325-364. [doi: 10.2139/ssrn.2206883]
33. Truong HM. Integrating learning styles and adaptive e-learning system: current developments, problems and opportunities. Comput Human Behav 2016 Feb;55:1185-1193. [doi: 10.1016/j.chb.2015.02.014]
34. Yang HD, Yoo Y. It's all about attitude: revisiting the technology acceptance model. Decis Support Syst 2004 Oct;38(1):19-31. [doi: 10.1016/s0167-9236(03)00062-9]
35. Topol E. The Creative Destruction Of Medicine: How The Digital Revolution Will Create Better Health Care. New York: Basic Books; 2013.
36. He J, Baxter SL, Xu J, Xu J, Zhou X, Zhang K. The practical implementation of artificial intelligence technologies in medicine. Nat Med 2019 Jan;25(1):30-36. [doi: 10.1038/s41591-018-0307-0] [Medline: 30617336]
37. Fraenkel J, Wallen N, Hyun H. How to Design and Evaluate Research in Education. New York, USA: McGraw-Hill; 2011.

Abbreviations

CCHS: Chainama College of Health Sciences
e-learning: electronic learning
e-platform: electronic health platform
IS: information system
ISSM: information system success model
MLP: medical licentiate practitioner
SY: study year
TAM: technology acceptance model

Edited by G Eysenbach; submitted 31.07.19; peer-reviewed by E Da Silva, E Micaiah, D Mazur; comments to author 17.08.19; revised version received 18.08.19; accepted 19.08.19; published 09.10.19.

Please cite as:
Barteit S, Neuhann F, Bärnighausen T, Bowa A, Wolter S, Siabwanta H, Jahn A
Technology Acceptance and Information System Success of a Mobile Electronic Platform for Nonphysician Clinical Students in Zambia: Prospective, Nonrandomized Intervention Study
J Med Internet Res 2019;21(10):e14748
URL: http://www.jmir.org/2019/10/e14748/
doi:10.2196/14748
PMID:31599731

Review

Methodological Challenges in Randomized Controlled Trials on Smartphone-Based Treatment in Psychiatry: Systematic Review

Morten Lindbjerg Tønning[1], MD; Lars Vedel Kessing[1], MD, Dr med; Jakob Eivind Bardram[2], MSc, PhD; Maria Faurholt-Jepsen[1], MD, Dr med

[1]Copenhagen Affective Disorder Research Center, Psychiatric Center Copenhagen, Rigshospitalet, University Hospital of Copenhagen, Copenhagen, Denmark

[2]Department of Health Technology, Technical University of Denmark, Lyngby, Denmark

Corresponding Author:
Morten Lindbjerg Tønning, MD
Copenhagen Affective Disorder Research Center
Psychiatric Center Copenhagen, Rigshospitalet
University Hospital of Copenhagen
Blegdamsvej 9
Copenhagen, 2100
Denmark
Phone: 45 28975755
Email: Morten.lindbjerg.toenning@regionh.dk

Abstract

Background: Smartphone-based technology is developing at high speed, and many apps offer potential new ways of monitoring and treating a range of psychiatric disorders and symptoms. However, the effects of most available apps have not been scientifically investigated. Within medicine, randomized controlled trials (RCTs) are the standard method for providing the evidence of effects. However, their rigidity and long time frame may contrast with the field of information technology research. Therefore, a systematic review of methodological challenges in designing and conducting RCTs within mobile health is needed.

Objective: This systematic review aimed to (1) identify and describe RCTs investigating the effect of smartphone-based treatment in adult patients with a psychiatric diagnosis, (2) discuss methodological challenges in designing and conducting individual trials, and (3) suggest recommendations for future trials.

Methods: A systematic search in English was conducted in PubMed, PsycINFO, and EMBASE up to August 12, 2019. The search terms were (1) psychiatric disorders in broad term and for specific disorders AND (2) smartphone or app AND (3) RCT. The Consolidated Standards of Reporting Trials electronic health guidelines were used as a template for data extraction. The focus was on trial design, method, and reporting. Only trials having sufficient information on diagnosis and acceptable diagnostic procedures, having a smartphone as a central part of treatment, and using an RCT design were included.

Results: A total of 27 trials comprising 3312 patients within a range of psychiatric diagnoses were included. Among them, 2 trials were concerning drug or alcohol abuse, 3 psychosis, 10 affective disorders, 9 anxiety and posttraumatic stress disorder, 1 eating disorder, and 1 attention-deficit/hyperactivity disorder. In addition, 1 trial used a cross-diagnostic design, 7 trials included patients with a clinical diagnosis that was subsequently assessed and validated by the researchers, and 11 trials had a sample size above 100. Generally, large between-trial heterogeneity and multiple approaches to patient recruitment, diagnostic procedures, trial design, comparator, outcome measures, and analyses were identified. Only 5 trials published a trial protocol. Furthermore, 1 trial provided information regarding technological updates, and only 18 trials reported on the conflicts of interest. No trial addressed the ethical aspects of using smartphones in treatment.

Conclusions: This first systematic review of the methodological challenges in designing and conducting RCTs investigating smartphone-based treatment in psychiatric patients suggests an increasing number of trials but with a lower quality compared with classic medical RCTs. Heterogeneity and methodological issues in individual trials limit the evidence. Methodological recommendations are presented.

(J Med Internet Res 2019;21(10):e15362) doi:10.2196/15362

KEYWORDS

psychiatry; methodology; smartphone; mHealth; mobile Health; digital health; digital psychiatry; systematic review

Introduction

Background

Psychiatric disorders represent a major burden of disease worldwide with a significant impact on the quality of life, socioeconomic factors, and life expectancy [1]. In 2010, the worldwide expenses because of mental illness were estimated to be between US $2.5 trillion and US $8.5 trillion [2]. Across European countries, 27% of the adult population suffers from at least one psychiatric disorder [3]. At the same time, there is a gap between the need for treatment and the number of patients receiving treatment. The number of patients who do not receive treatment for their disorder is 35% to 50% in high-income countries and 76% to 85% and in low- and middle-income countries [4].

In 2011, the World Health Organization stated that "the use of mobile and wireless technologies to support the achievement of health objectives (mHealth) has the potential to transform the face of health service delivery across the globe" [5]. The number of smartphone users exceeded 2.5 billion people in 2018 [6], and in high-income countries, 80% of the population own and use a smartphone [7].

Smartphones are a promising tool in the field of psychiatry. They are widely used and always at hand, allowing for delivery of treatment to patients in real-time and naturalistic settings, and can augment already available treatments. At the same time, smartphones contain several sensors and technologies enabling patients, researchers, and clinicians to access information about physical and social activities [8].

Mobile health (mHealth) and especially smartphone-based technology and solutions are developing at an enormous speed, driven mainly by software and computer scientists and private companies. Thus, most available apps have not been scientifically investigated, and the validity, treatment effect, and safety have been sparingly investigated [9,10]. Nevertheless, hundreds of apps claiming to help or monitor psychiatric disorders are already available in app stores [11].

Evidence, Randomized Controlled Trials, and Interdisciplinary Research

Randomized controlled trials (RCTs) represent the methodological golden standard of excellence in medical research for the investigation of possible positive and negative effects of treatment interventions [12]. The importance of RCTs in medical science is mainly because of their ability to eliminate confounding of known and unknown nature. If properly designed and conducted, RCTs are especially useful for examination of small or moderate positive and negative effects [12].

The design and conduct of RCTs within smartphone-based treatment interventions and other electronic treatments should follow the RCT standards within medicine, while also taking into account particular challenges with electronic effect research. In RCTs, in general, there is a significant time gap from design and trial initiation to the publication of results. A recent Australian study of publicly funded clinical trials showed a median time of 7.1 years (95% CI 6.3-7.6 years) from funding to the main article on trial results being published [13]. This time gap is particularly problematic when testing smartphone-based treatment interventions as the technological development taking place within the timespan of an RCT is enormous. The technology tested is at risk of being outdated when results are published and before being taken into clinical usage [14]. Furthermore, the locked nature of treatment interventions tested in RCTs contrasts with the constantly evolving and iterative nature of app solutions [15].

Previous Reviews

Other groups have previously suggested changes to the development process, to speed up the process from idea to publication [14,16,17]. A previous review from 2013 [18] identified effect studies of smartphone-based treatments within psychiatry (including stress). This review was not limited to include RCTs, and thus, it also included pre- to posttest studies. Only 8 studies were identified despite an extensive search and broad inclusion criteria.

Symptom-specific reviews within branches of psychiatry have been done as well. As an example, Firth et al [19] found studies regarding depressive *symptoms* and conducted a meta-analysis on the possible effects of smartphone-based treatments on these symptoms. A total of 18 RCTs were identified; however, often, a psychiatric diagnosis was not present as the focus was on depressive symptoms. This study focused on RCTs only. The focus was on research methods and design in relation to smartphone-based treatment interventions within psychiatry. This study was limited to RCTs providing information on diagnostic measures to ensure that participants suffers from a psychiatric diagnosis, excluding trials focusing on symptoms in healthy populations.

Aims

This systematic review aimed to identify and describe all available RCTs using smartphones as (part of) a treatment intervention conducted in the field of psychiatry using sufficient diagnostic measures. Furthermore, it aimed to describe the methodology of these individual trials and discuss methodological challenges related to designing and conducting RCTs within smartphone-based treatment in psychiatric disorders using the Consolidated Standards of Reporting Trials (CONSORT) electronic health (eHealth) checklist [20] as a reference and to provide recommendations for future trials within the area.

Methods

Methods of the review and eligibility criteria were established in advance by 3 of the authors (MLT, MFJ, and LVK). Minor modifications such as adding further information to be extracted were made to the review protocol during the review process.

Trial Selection

Original trials reporting on smartphone-based treatment interventions investigated in RCTs, including adult patients with psychiatric disorders, were eligible for review. In addition, peer-reviewed articles, posters, and conference abstracts were eligible for review. If multiple articles reported on overlapping populations, the article presenting the largest population was included. No restrictions regarding sample size were applied.

The exclusion criteria were as follows: (1) children younger than 18 years; (2) psychiatric symptoms as part of somatic disorders (ie, preoperation anxiety or depressive symptoms in terminal cancer patients); (3) trials concerning stress, cigarette smoking, low intelligence quotient (IQ), and isolated sleep problems without psychiatric disorders; (4) trials with individuals who self-identified as having a psychiatric diagnosis but without diagnostic reassurance; (5) trials reporting on symptoms without diagnoses (ie, depressive symptoms or alcohol usage among college students); (6) trials using internet therapy without an active smartphone-based component (ie, if the Web page was accessible from a smartphone browser); (7) trials using only cell phones in traditional ways with text messages and phone calls (not using smartphone-based features); (8) trials using smartphones as a screen for virtual-reality setups as the primary treatment component; and (9) trials not available in English.

Information Source, Trial Selection, and Data Extraction

We conducted a systematic search covering PubMed, PsycINFO, and EMBASE on April 23, 2018, and it was last updated on August 12, 2019. Only articles from 2008 onwards were included (the time of the release of the first smartphone). References from articles and other reviews were also examined, but they did not result in any additional trials to include. The trial selection was conducted by 2 researchers (MFJ and MLT), and articles with doubt about the relevance were discussed between the 2 of them. Full-text articles for possible relevant trials were obtained if the abstract and title did not supply sufficient information. The search strategy included (1) psychiatric disorder as a broad term and for specific diseases AND (2) smartphone or app AND (3) RCT. A wide variety of text words were used to include trials published within the last 6 months that had not yet been indexed with Medical Subject Headings terms. Search strategy in PubMed was as follows:

> *(((Smartphone[MeSH terms] OR mobile application[MeSH Terms] OR smartphone OR mobile application OR smart phone OR mobile phone OR app OR apps OR cell phone OR Iphone* OR IOS OR Android phone OR smartphones OR mobile applications OR smart phones OR mobile phones OR cell phones)) AND ((((((((mental disorder[MeSH Terms]) OR (mental disorder OR mental disorders OR mental disease OR mental diseases OR mental diagnose OR psychiatric disease OR psychiatric diseases OR psychiatric disorders OR psychiatric disorder OR psychiatric diagnose)) OR ((drug OR substance OR prescription drug OR alcohol OR narcotic OR heroin OR amphetamine OR cocaine OR marijuana OR opioid OR morphine OR phencyclidine) AND (abuse OR dependence OR addiction))) OR (feeding disorder OR feeding disorders OR eating disorders OR eating disorder OR anorexia OR bulimia OR binge eating)) OR (autism OR autistic OR Asperger disease OR Aspergers disease) OR Asperger disorder OR Aspergers disorder OR ADHD OR attention deficit disorder OR ADD OR attention deficit hyperactivity disorder)) OR (personality disorder OR personality disorders OR obsessive-compulsive personality OR compulsive personality OR obsessive personality OR psychopath OR sociopathic OR antisocial OR passive-dependent personality OR dyssocial OR schizoid OR schizotypal)) OR (schizophrenia OR psychoses OR psychosis OR psychotic OR paranoid OR schizoaffective OR schizophreniform OR delusional)) OR (major depressive disorder OR unipolar depression OR unipolar disorder OR depressive syndrome OR endogenous depression OR neurotic depression OR melancholia OR cyclothymic OR dysthymic OR mood disorder OR mood disorders OR affective disorder OR affective disorders OR bipolar OR manic-depressive OR mania OR manic) OR (anxiety OR anxieties OR panic disorder OR agoraphobia OR obsessive disorder OR compulsive disorder OR obsessive-compulsive disorder OR phobic disorder OR phobic disorders OR PTSD OR posttraumatic stress disorder OR post-traumatic stress disorder OR post traumatic stress disorder))) AND ((randomized controlled trial[MeSH Terms]) OR (randomized controlled trial OR randomised controlled trial OR randomised OR randomized OR RCT OR randomized clinical trial OR randomiced clinical trial OR randomized clinical trial OR randomized controlled clinical trial OR randomised controlled clinical trial))*

Data were extracted by using the CONSORT eHealth checklist as a template for data extraction [20]. The following data were extracted:

- Author, year, country, trial design, trial registration, protocol availability, patients' age and gender, sample size and use of power calculations, length of treatment intervention, and follow-up period.
- Recruitment and diagnostics procedures of patients, recruitment length, outcome measures, well-defined hierarchy of outcome measures, and collection of outcome data.
- Description of treatment intervention and comparator, use of blended treatment and standard treatments, affiliation with industry and technology descriptions.
- Title according to CONSORT recommendations; use of prompts, platform choice, and possible lent-out of smartphones; economic compensation; use of placebo smartphones; methodological information regarding randomization and blinding procedures; information about the statistical approach to technical updates and whether updates and technical crashes or down periods were

reported; possible harms; adherence to the smartphone system; and baseline data on patients' technological skills and funding information.

A data extraction template is provided in the supplementary material (Multimedia Appendix 1). The extracted data are presented in 4 tables. Data were initially retrieved by one author (MLT) and subsequently and independently by another author (MFJ). Any disagreements were solved between MLT and MFJ.

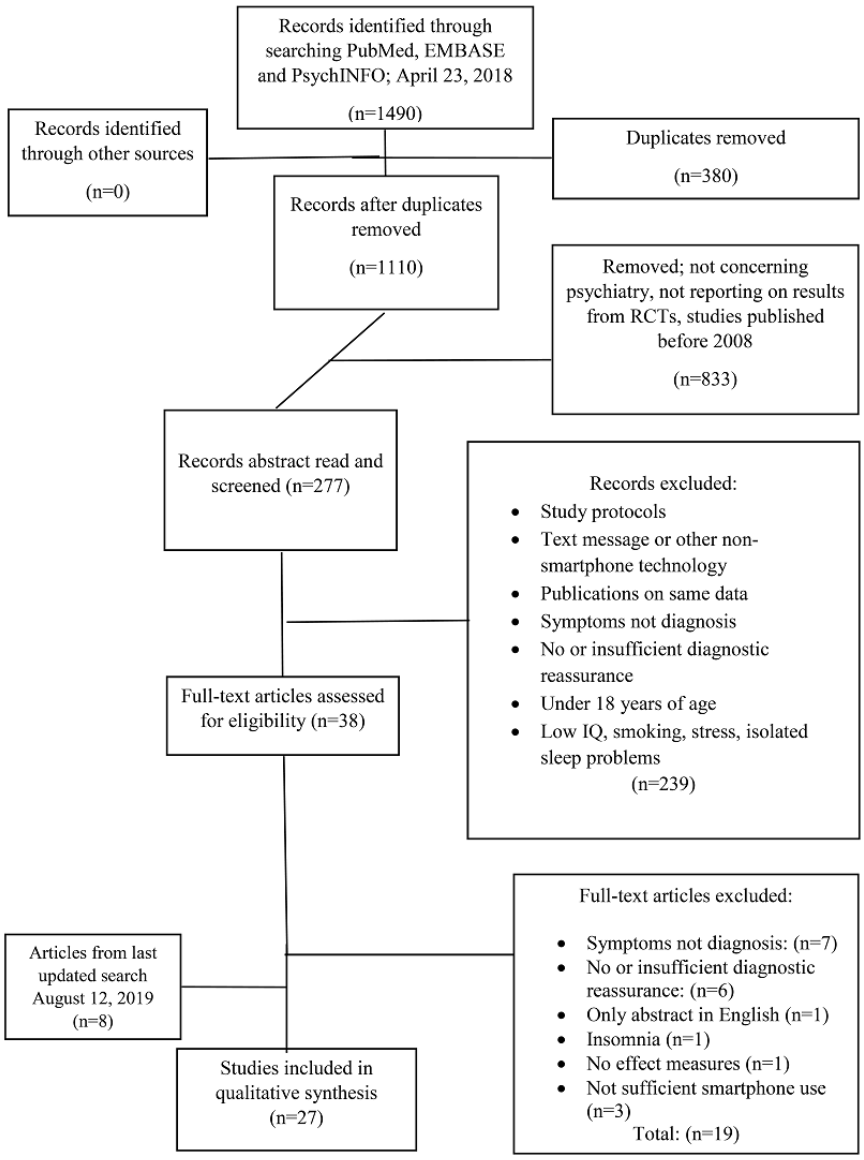
Three tables describe the trials individually, with various focus. The fourth uses the CONSORT eHealth checklist [20] as a template to summarize the relevant findings according to these in a systematic way and hereby partly summarizes the relevant findings from previous tables, but it also includes other new relevant information. The tables are presented in relevant sections in the result-section.

Results

Trial Selection

Figure 1 presents a flow diagram represented according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines [21], showing the results of the literature search and selection of trials. The initial literature search in PubMed, EMBASE, and PsycINFO on April 23, 2018, resulted in a total of 1490 articles. Of these, 380 duplicates were removed, resulting in 1110 articles. Furthermore, 833 articles were excluded based on the title and year of publication with the main reasons for exclusion being as follows: not concerning psychiatric disorders (eg, HIV, Alzheimer disease, and obesity), not reporting on results from RCTs, and published before 2008. This led to a total of 277 remaining articles, from which abstracts were examined.

Figure 1. Preferred Reporting Items for Systematic Reviews and Meta-Analyses flow diagram displaying information on article flow from initial search to final inclusion. IQ: intelligence quotient; RCT: randomized controlled trial.



Out of these 277 articles, 239 were excluded based on abstracts, with majority of reasons being as follows: trial protocols, text message or other nonsmartphone technology, publications on same data (the publication with the most data was included), measuring on symptoms—not diagnosis, no or insufficient diagnostic reassurance, participants younger than 18 years, topics concerning low IQ, smoking, stress, or isolated sleep problems.

This led to 38 articles that were printed for full-text reading. Among them, 19 articles were subsequently excluded because of the following: reporting symptoms, not disorders (n=7); no or insufficient diagnostic reassurance (n=6); only abstract in English (n=1); insomnia (n=1); no outcome effect measures (n=1); and not using smartphone as part of treatment (ie, only using smartphones to reach the patient or allowing patients to answer emails and questionnaires on smartphone; n=3). References from articles and other reviews were also examined, but they did not result in any additional trials to include, giving a total of 19 eligible articles from the initial search.

The search was updated on August 12, 2019, resulting in 8 new articles. Thus, a total of 27 unique trials [22-48] were identified and included in this review, including 5 with diagnostic estimates based on questionnaires using relevant cutoff scores [44-48]. Included trials are described individually in the tables.

Trial Characteristics

Across all included trials, a total of 3,312 patients were included. The included trials represented a range of psychiatric disorders. Numbers in parentheses represent relevant chapter and coding according to International Statistical Classification of Diseases and Related Health Problems-10 (ICD-10): 2 on drug or alcohol misuse (F10-F19) [24,37], 3 on psychosis (F20-29) [35,38,41], 10 on affective disorders (F30-39) [22,25-28,32,36,40,42,45] (comprising 3 on bipolar disorder and 7 on unipolar depressive disorder), 9 on anxiety and posttraumatic stress disorder posttraumatic stress disorder (F40-F48) [23,30,34,39,43,44,46-48], 1 on eating disorders (F50-F59) [31], and 1 on attention-deficit disorder (F90-F98) [29]. One trial included patients with severe mental disorders across ICD-10 diagnoses [33].

Most trials had an overrepresentation of females, except trials concerning veterans, schizophrenia, or substance or alcohol abuse. Patients' ages ranged from 18 to 73 years. In the majority of the trials, the average age was 40 years.

Trial Design and Reporting

The number of RCTs testing smartphone solutions in patients with a psychiatric disorder increased with time, especially from 2018 onward. Some trials [24,27,29-33,35,40] were leaning toward traditional RCT designs conducted in clinical settings and using the CONSORT checklist for reporting and designing trials [20]. Some other trials did not provide clear information regarding these issues, leaving the reader with lacking information on the design and conduct of the trial.

Overall, 11 trials had a sample size above 100 patients [24,30,32-34,39-41,44,45,48]. One trial [40] had been repeated using the same intervention, trial design, and outcome measures as a previous trial [27].

A total of 14 articles included information about trial registration, such as the clinicaltrials.gov [22,24-27,29-33,35,39-41]. For the remaining trials (including the study by Krzystanek et al [41], where we could not find any information on the registration information provided) it was not possible to see whether the primary outcome measure was predefined in the primary hypothesis leading to the trial design. In addition, 5 trials referred to a published trial protocol in the article [27,30,32,35,40]. One additional article attached the study protocol in the supplementary section when published [28].

Information about the individual trial designs, registration, sample characteristics, and length of intervention and follow-up is presented in Table 1.

Table 1. Randomized controlled trials involving smartphones in the field of psychiatry identified in systematic search updated in August 2019. Description of basic information and trial design. The bottom 5 trials in *italics* are trials with diagnoses solely based on questionnaires.

Author, year of publication	Country	Trial design	Protocol[a]	Female/total	Analyzed (power calculation)	Age (years), mean (SD)	Intervention length (weeks)	Posttreatment follow-up (weeks)
Watts et al, 2013 [22]	Australia	Pilot RCT[b]	N/A[c]	28/35	25 (N/A)	41 (12.4)	8	12
Dagöö et al, 2014 [23]	Sweden	RCT	N/A	39/57	52 (N/A)	52	9	12
Gustafson et al, 2014 [24]	United States	RCT	N/A[c]	137/349	349 (350)	38 (10)	32	16
Ly et al, 2014 [25]	Sweden	RCT	N/A[c]	57/81	81 (N/A)	36.0 (10.8)	8	16
Depp et al, 2015 [26]	United States	RCT	N/A[c]	48/82	82 (N/A)	47.5	10	14
Faurholt-Jepsen et al, 2015 [27]	Denmark	RCT	Published[c]	45/67	78 (56)	18-60[d]	24	No
Ly et al, 2015 [28]	Sweden	NI[e] RCT	Attached	65/95	93 (93)	18-73[d]	9	24
Moëll et al, 2015 [29]	Sweden	RCT	N/A[c]	39/57	57 (N/A)	36.8 (10.9)	6	No
Ivanova et al, 2016 [30]	Sweden	RCT[f]	Published[c]	98/152	152 (150)	35.3	10	52
Hildebrandt et al, 2017 [31]	United States	RCT	N/A[c]	55/66	66 (80)	32.1	12	24
Mantani et al, 2017 [32]	Japan	RCT	Published[c]	87/164	164 (164)	25-59[d]	8	17
Ben-Zeev et al, 2018 [33]	United States	RCT	N/A[c]	67/163	163 (160)	49 (10)	12	12
Boettcher et al, 2018 [34]	Sweden/Germany	RCT[f]	N/A	161/209	209 (N/A)	35.4 (12.4)	12	52
Bucci et al, 2018 [35]	England	Pilot RCT	Published[c]	18/36	36 (N/A)	N/A	12	10
Hur et al, 2018 [36]	Republic of Korea	Pilot RCT	N/A	26/48	34 (N/A)	≈24	3	No
Liang et al, 2018 [37]	China	Pilot RCT	N/A	21/74	75 (N/A)	41.6 (8.0)	4	No
Schlosser et al, 2018 [38]	United States	RCT	N/A	15/43	43 (N/A)	24	12	12
Stolz et al, 2018 [39]	Switzerland	RCT[f]	N/A[c]	94/150	150 (141)	35	12	12
Faurholt-Jepsen et al, 2019 [40]	Denmark	RCT	Published[c]	76/129	129 (117)	43 (12.4)	36	No
Krzystanek et al, 2019 [41]	Poland	RCT	N/A[c]	116/290	Varying	32.1 (6.2)	52	No
Stiles-Shields et al, 2019 [42]	United States	Pilot RCT	N/A	N/A/30	27 (N/A)	N/A	6	4
Teng et al, 2019 [43]	Taiwan	RCT[f]	N/A	61/82	82 (N/A)	≈21.5	4	4
Enock et al, 2014 [44]	United States	RCT	N/A	205/429	326[g] (N/A)	18-68[d]	4	8
Roepke et al, 2015 [45]	United States	RCT[f]	N/A	197/283	283 (207)	40.2 (12)	4	2
Miner et al, 2016 [46]	United States	Pilot RCT	N/A	40/49	49 (N/A)	45.7 (13.9)	4	4

Author, year of publication	Country	Trial design	Protocol[a]	Female/total	Analyzed (power calculation)	Age (years), mean (SD)	Intervention length (weeks)	Posttreatment follow-up (weeks)
Possemato et al, 2015 [47]	United States	Pilot RCT	N/A	1/20	20 (N/A)	42	8	No
Kuhn et al, 2017 [48]	United States	RCT	N/A	83/120	120 (120)	39	12	12

[a]Mentioned in the article.

[b]RCT: randomized controlled trial.

[c]Trial is registered.

[d]Age interval (mean not given in the article).

[e]NI: noninferiority.

[f]The trial had 3 arms.

[g]Included in the analysis if score is greater than cutoff.

Settings and Diagnostic Procedures

A total of 8 trials were conducted in traditional clinical settings, with patients referred by clinical staff or contacted in their treatment clinic [24,27,32,33,35,37,40,41]. Others used remote designs where patients who self-identified as having a psychiatric disorder applied for participation [22,23,25,26,28,30,34,38,39,42-44,46,48]. None of the trials compared participants with nonparticipants. All but 2 trials [43,44] presented a flowchart of eligible subjects with varying details on reasons not to participate. Trials with open recruitment (eg, internet forums, Web pages, Facebook, and advertising) [22,23,25,26,28-30,34,36,38,39,42-46,48] only had information on people who signed up. Completion rates were reported very differently and varied from 163 patients out of 164 (99.4% [163/164]) completing the primary outcome [32] to 74 out of 283 patients (26.1% [74/283]) [45]. In addition, 12 trials compensated for participation in trial assessments with money or gift cards [26,33-38,42,43,46-48].

Validation and certainty of diagnosis varied substantially, and often, a pragmatic setup was used, leaving the validity of the obtained diagnosis with uncertainty.

A total of 5 trials based their diagnosis solely on clinical-based information [24,33,35,37,41]. Furthermore, 15 trials used research-based diagnoses without clinical information, using either questionnaires [44-46,48], phone calls [22,23,25,28,34,42], video interviews [38], or personal interviews [31,36,39,43] to validate the self-reported diagnoses. In addition, 7 trials included patients with a clinical diagnosis that was subsequently assessed and validated by the researchers using questionnaires [47], phone calls [29,30], or personal interviews [26,27,32,40].

Hypotheses and Use of Predefined Hierarchical Outcome Measures

Overall, 18 trials included clearly described hypotheses [22,24-28,30-32,34,36,38-40,44,45,47,48], and 8 trials [22,37,41-44,46,47] did not distinguish between primary and secondary outcome measures in the article and had no hierarchy of outcomes.

Outcome Measures

A total of 8 trials included objective or observer-based primary outcome measures: One trial tested the levels of drug use in the urine as a specific primary outcome measure [37]. Another trial used objective measurements of feasibility, use, and attrition as the primary outcome [35]. One trial tested a specific task of motivated behavior [38]. In addition, 4 trials used scores on clinical ratings as the primary outcome measure assessing the level of depression and mania [26,27,40], levels of bulimic episodes [31], or psychotic symptoms [41]. The remaining trials used patient-reported outcome measures.

Within trials of similar diagnoses, authors used different scales and measures for the primary outcome measure. For instance, in measuring depression scores, the following outcome measures were used as the primary outcome: Patient Health Questionnaire, Beck Depression Inventory, Hamilton Rating Scale for Depression 17 item, Montgomery–Åsberg Depression Rating Scale, and Center for Epidemiologic Studies Depression Scale [22,25-28,32,33,36,40,42,45].

Furthermore, 2 trials, with apps for training in attention bias modification, claimed to be double blinded, with no further explanation on how blinding was ensured [43,44]. One trial blinded app allocation for patients when testing 2 different apps [45]. The remaining trials did not blind patients for the intervention (active vs control). In 12 trials, the authors explicitly stated that they had used blinded assessments for outcome measures [25-29,31-33,35,38,40,44]. Within these 12 trials, 5 trials used the patient-reported outcome as the primary outcome measure in nonblinded patients [25,28,29,32,33]. Although blinded assessors collected these data, answers to the questionnaires were self-reported. One trial tested for the success of blinding [32].

Information about recruitment, diagnosis, outcome measures, and how these were obtained can be found in Table 2.

Table 2. Randomized controlled trials involving smartphones in the field of psychiatry identified in systematic search updated in August 2019. Description of participant characteristics and outcome data. The bottom 5 trials in *italics* indicate trials with diagnoses solely based on questionnaires.

Author, year of publication	Diagnosis	How was the diagnosis obtained?	Recruitment: *open/closed (recruitment length in months);* information	Primary outcome[a]	Questionnaire data collection
Watts et al, 2013 [22]	Major depressive disorder (R[b])	MINI[c] phone interview + PHQ-9[d]	*Open (3 months);* advertising + application on a Web page	*Questionnaire*(PHQ-9, BDI-II[e], K-10[f], and other)	N/A[g]
Dagöö et al, 2014 [23]	Social anxiety disorder (R)	MINI phone interviews SCID[h]	*Open (2011 and 2012);* advertising in media and Facebook.	*Questionnaire* (LSAS-SR[i])[a]	Internet platform
Gustafson et al, 2014 [24]	Alcohol use disorder (C[j])	From treatment centers (DSM-IV[k])	*Closed (17 months)*; from 3 nonprofit residential treatment centers	Risky drinking days[a]	Phone interview
Ly et al, 2014 [25]	Major depressive disorder (R)	MINI phone interview	*Open (N/A);* advertising in national media	*Questionnaires* (PHQ-9 and BDI-II)[a]	Internet platform
Depp et al, 2015 [26]	Bipolar disorder (R+C)	Medical records + MINI interview	*Open (N/A);* online, self-help groups, outpatient clinics, and communities	*Clinical ratings* (MADRS[l] and YMRS[m])[a]	N/A
Faurholt-Jepsen et al, 2015 [27]	Bipolar disorder (R+C)	From outpatient clinic + SCAN[n]	*Closed (18 months);* recruitment from specialized hospital function	*Clinical rating* (Hamilton Rating Scale for Depression 17 item and YMRS)[a]	Paper
Ly et al, 2015 [28]	Major depressive disorder (R)	MINI phone interview + PHQ-9	*Open (N/A);* advertising in media (for patients' self-identifying as depressed)	*Questionnaires* (BDI-II)[a]	Internet platform
Moëll et al, 2015 [29]	Attention-deficit/hyperactivity disorder. (R+C)	Medical records + DSM-IV phone	*Open (N/A);* by patient websites and Facebook	*Questionnaire* Adult Self-Reported Scale subscale[a]	Internet platform
Ivanova et al, 2016 [30]	Social anxiety/panic disorder (R+C)	MINI phone + questionnaires	*Open (2 months);* advertising in national medias	*Questionnaire* (Generalized Anxiety Disorder 7-item, LSAS-SR, + more a)[a]	Internet platform
Hildebrandt et al, 2017 [31]	Binge eating and bulimia (R)	SCID interview + questionnaires	*Both (N/A);* community advertising and referrals	*Clinical ratings* (objective bulimic episodes Eating Disorder Examination)[a]	Paper and in-app
Mantani et al, 2017 [32]	Major depressive disorder (R+C)	Personal by treating physician	*Closed (25 months);* recruited by treating physicians	*Telephone* (blinded; PHQ-9)[a]	Telephone assessment
Ben-Zeev et al, 2018 [33]	Severe mental illness[o] (C)	Chart diagnosis	*Closed (27 months);* identified from medical records, recruited by clinical teams	*Questionnaire* (Symptom Checklist-9) Engagement (*objective*)[a]	N/A
Boettcher et al, 2018 [34]	Social anxiety disorder (R)	DSM-IV (phone)	*Open (N/A);* advertising in national media, Facebook	*Questionnaire* (LSAS-SR)[a]	Internet platform
Bucci et al, 2018 [35]	Early psychosis (C)	From outpatient clinic	*Closed (7 months);* From early intervention for psychosis service	*Objective* measurements of feasibility and attrition[a]	N/A
Hur et al, 2018 [36]	Depression (other)[p] (R)	SCID non-patient interview + questionnaires	*Open (N/A);* advertising, online recruitment, posters and clinic	*Questionnaires* (Dysfunctional Attitude Scale)[a]	N/A
Liang et al, 2018 [37]	Substance use disorder (C)	From methadone treatment clinics	*Closed (11 months);* from methadone maintenance clinics and via social workers	*Robust objective measure* (drug use measured in urine)	Interviews
Schlosser et al, 2018 [38]	Schizophrenia spectrum disorders (R)	DSM-IV video-interview	*Open (N/A);* Craigslist, advertising, and information boards	Motivated behavior measured by trust task (objective task)[a]	Internet platform
Stolz et al, 2018 [39]	Social anxiety disorder (R)	DSM-IV (master students)	*Open (N/A);* media and online forums	*Questionnaires* (Social Phobia Scale, SIAS[q], LSAS-SR)[a]	Internet platform

Author, year of publication	Diagnosis	How was the diagnosis obtained?	Recruitment: *open/closed (recruitment length in months);* information	Primary outcome[a]	Questionnaire data collection
Faurholt-Jepsen et al, 2019 [40]	Bipolar disorder (R+C)	SCAN	*Closed (N/A);* patients previously treated in specialized function invited	*Clinical ratings* (HDRS-17 + YMRS)[a]	Paper
Krzystanek et al, 2019 [41]	Paranoid schizophrenia (C)	N/A (enrolled from treatment centers)	*Closed (7 month);* enrolled from 27 treatment centers	Many outcomes; clinical ratings (video; eg, Positive and Negative Syndrome Scale)	N/A
Stiles-Shields et al, 2019 [42]	Depression (R)	Quick Inventory of Depressive Symptomatology + MINI phone interview	*Open (5 month);* Craigslist	*Questionnaires* (eg, PHQ and usability)	Internet platform
Teng et al, 2019 [43]	Generalized Anxiety Disorder (R)	Questionnaire + DSM-IV subscale	N/A	*Questionnaires* (State-Trait Anxiety Inventory, BDI-II + more)	N/A
Enock et al, 2014 [44]	Social anxiety disorder (R)	Questionnaire with cutoff	*Open (40 months);* news articles, message boards, craigslist.org, Google	*Questionnaire* (LSAS-SR, SIAS-17, and other)	Internet platform
Roepke et al, 2015 [45]	Depression (R)	Questionnaire CES-D[r] above 16	*Open (5 months);* Website and Craigslist	*Questionnaire* (CES-D at posttest)[a]	Internet platform
Miner et al, 2016 [46]	PTSD[s] (R)	Questionnaire PCL-C[t] >30	*Open (16 months);* flyers, websites, Craigslist	*Questionnaire* (PCL-C) + feasibility and acceptability	Internet platform
Possemato et al, 2016 [47]	PTSD (R+C)	Screened for PTSD + PCL >40	*Closed (4 months);* from veteran care unit, screened for PTSD.	Feasibility metrics + *Questionnaires* (PCL-S[u], PHQ-9)	N/A
Kuhn et al, 2017 [48]	PTSD (R)	Questionnaires PCL-C>34	*Open (10 months);* flyers, media, social media, Craigslist	*Questionnaire* (PCL–C)[a]	Internet platform

[a]Well-defined hierarchy in outcome measures.

[b]R: research based.

[c]MINI: Mini International Neuropsychiatric Interview.

[d]PHQ-9: Patient Health Questionnaire-9.

[e]BDI-II: Beck Depression Inventory.

[f]K-10: Kessler Psychological Distress Scale.

[g]N/A: not applicable.

[h]SCID: Structured Clinical Interview.

[i]LSAS-SR: Liebowitz Social Anxiety Scale—self-reported.

[j]C: clinical based.

[k]DSM: Diagnostic and Statistical Manual of Mental Disorders.

[l]MADRS: Montgomery–Åsberg Depression Rating Scale.

[m]YMRS: Young Mania Rating Scale.

[n]SCAN: Schedules for Clinical Assessment in Neuropsychiatry.

[o]Schizophrenia, schizoaffective disorder, bipolar disorder, and major depressive disorder.

[p]Other specified depressive disorder.

[q]SIAS: Social Interaction Anxiety Scale.

[r]CES-D: Center for Epidemiologic Studies Depression Scale.

[s]PTSD: posttraumatic stress disorder.

[t]PCL-C: Post-Traumatic Checklist—Civilian.

[u]PCL-S: Post-Traumatic Checklist Scale.

Interventions

Although the included trials shared the concept of smartphones as a core feature of their interventions, the interventions were very different and tested in diverse settings with heterogeneous patient groups. Generally, 8 trials tested the app as a stand-alone treatment [35-37,44-48]. The remaining trials used variations of blended treatment and clinical support. Overall, 18 trials used prompts to engage users either from the app or by the investigators [22,26-28,30,32-44].

A glimpse of the diversity of interventions is presented here. Some trials used smartphones as a part of internet-based therapy, whereas others used smartphones as augmentation for face-to-face therapy or standard treatment programs. Most trials developed specific apps, whereas others used either commercially available or previously developed apps. Some interventions were interactive, whereas others had more static content. Some of the interventions made use of the unique possibilities a smartphone represents such as using global positioning system or allowing to interact with peers, relatives, or professionals, whereas others resemble classic internet therapies made for small smartphone screens to be more convenient and accessible to patients.

Control Group

One of the included trials involved the use of a placebo app, consisting of an inactive version of the software with limitations, but with no further description [41]. In addition, 2 trials (that both tested the effect of attention bias modification via a dot probe on the smartphone screen) used a different version of the app for the control group, setting the dot appearance at random instead of following a predefined pattern that was central to the treatment [43,44]. In 1 trial on bipolar disorder, the participants in the control group received a *placebo* smartphone without the app system [27]. No other trials mentioned attempts on placebo treatment. Overall, 5 trials used standard treatment as the comparator [24,27,35,38,40], and 11 trials used a waitlist control group [29,30,34,38,39,41,42,44-46,48]. Furthermore, 4 trials used some sorts of clinical intervention as the comparator [26,28,31,33], and 5 trials compared the intervention with another app [23,25,35,36,42].

Further description of the smartphone-based treatment interventions, availability of technology, and author affiliation with industry can be found in Table 3 and is summarized in Table 4.

Table 3. List of randomized controlled trials involving smartphones in the field of psychiatry identified in systematic search updated in August 2019. Description of intervention and control group as well as authors cooperation with the industry. The bottom 5 trials in *italics* indicate trials with diagnoses solely based on questionnaires.

Author, year of publication	Short description of the intervention and main components. If available, the app name is displayed in italics.	Comparator. treatment received by the control groups	Blended treatment (BT)/app alone (AA)	TAU[a]	Cooperation/affiliation with the industry[b]	Description of technology available for the reader
Watts et al, 2013 [22]	CBT[c]-based "get happy program" with comic book–like lessons + homework activities	PC version of the same program	BT (limited clinician contact)	N/A[d]	No	Brief description and few screenshots
Dagöö et al, 2014 [23]	Guided internet-based CBT adapted for mobile phone administration	Another app similar therapist contact	BT (limited clinician contact)	No	N/A	Brief descriptions
Gustafson et al, 2014 [24]	*A-chess:* app with static self-help content and interactive features with therapist feedback	TAU	BT	Yes	No	App fully available online and a description of the app is attached
Ly et al, 2014 [25]	App delivering behavioral activation psychotherapy with possible but limited clinician contact	Mindfulness app, similar therapist contact	BT (limited clinician contact)	No	The first author has a similar app on the open market	Good descriptions and screenshots
Depp et al, 2015 [26]	*PRISM:* interactive monitoring and intervention linking mood and activities with self-management strategies	Active control monitoring on paper	BT	N/A	No	Thorough descriptions but no technical reports or screenshots
Faurholt-Jepsen et al, 2015 [27]	*MONARCA:* self-monitoring with a double feedback loop between clinic (nurse) and patient	TAU + nurse contact + phone without app	BT	Yes	No	Thorough descriptions and screenshots in the protocol
Ly et al, 2015 [28]	Four therapy sessions and a smartphone app, based on behavioral activation, used between sessions	Full behavioral activation (10 sessions)	BT	No	N/A	Brief descriptions and screenshots
Moëll et al, 2015 [29]	*Living smart:* Guided online course to structure life using multiple already available apps	Waitlist control	BT	No	N/A	Multiple already-available apps
Ivanova et al, 2016 [30]	Internet therapy + an app[e] promoting change corresponding to the core treatment program, with therapist support	Waitlist or intervention without therapist support	BT	No	2 authors employed by a technology company; 1 developed a similar app	Description and screenshots available in the protocol
Hildebrandt et al, 2017 [31]	*NOON self-monitoring:* App as an augmentation of traditional guided self-help	Guided self-help therapy without an app	BT	No	3/5 authors have a connection to NOOM who developed the app	Short descriptions and no screenshots
Mantani et al, 2017 [32]	*Kokoro:* CBT-based self-help app with 8 sessions presented by cartoons + fixed-dose medicine shift	Medicine shift with fixed dose and no app	BT[f]	No	2 of the authors developed the app	A thorough report describing the app in detail
Ben-Zeev et al, 2018 [33]	*FOCUS:* Multimodal smartphone intervention including self-assessments and on-demand functions	Clinic-based group intervention	BT	No	First author had a consulting agreement with technology company	A short description in the text - further in supplement
Boettcher et al, 2018 [34]	*Challenger:* App promoting exposure exercise through interactive challenges + internet-based therapy	2) Waitlist control or 3) internet therapy alone	BT	No	The third author founded the app company	A short description + referral to further information
Bucci et al, 2018 [35]	*Actissist:* Self-help app that asks questions and has automated responses and various static supportive content	TAU + another app	AA	Yes	N/A	Descriptions of app and screenshots

Author, year of publication	Short description of the intervention and main components. If available, the app name is displayed in italics.	Comparator. treatment received by the control groups	Blended treatment (BT)/app alone (AA)	TAU[a]	Cooperation/affiliation with the industry[b]	Description of technology available for the reader
Hur et al, 2018 [36]	*TODAC:* A scenario-based CBT app to reduce dysfunctional beliefs	Mood diary app	AA	N/A	No	Descriptions of the app modules, and small screenshots
Liang et al, 2018 [37]	*S-health:* Simple smartphone app that sends messages, controls cravings, and has a survey	Receiving text messages about various topics	AA	Yes	N/A	Survey and screenshots available + short description
Schlosser et al, 2018 [38]	*Prime:* Personalized real-time intervention for motivational enhancement. App-based online community	TAU/waitlist control	BT	Yes	No	Short description, but no screenshots
Stolz et al, 2018 [39]	Mobile version of validated psychoeducative self-help program with 8 modules based on cognitive therapy	Waitlist control or PC version	BT (limited clinician contact)	No	N/A	Short description of modules, but no screenshots
Faurholt-Jepsen et al, 2019 [40]	*MONSENSO:* Self-monitoring + objective monitoring with a double feedback loop between nurse and patient	TAU + offer to borrow a smartphone	BT	Yes	2 coauthors are shareholders in Monsenso	Thorough descriptions and screenshots in the protocol
Krzystanek et al, 2019 [41]	*MONEO:* Medication reminder, cognitive training, information bank, and "tele visits" with the investigator	Inactive version + monthly video examination	BT	N/A	No	Refers to online supplementary that was not possible to find
Stiles-Shields et al, 2019 [42]	*Boost Me* (behavioral app) or *Thought Challenger* (a cognitive app): Both with brief weekly coaching	2 different apps and 1 waitlist control	BT (with limited coaching)	No	Last author has an ownership interest in Actualize Therapy	Both apps are available free online
Teng et al, 2019 [43]	Home-delivered attention bias modification training with dot probe on screen	Control group with random dot or waitlist	BT	N/A	No	Short descriptions and few screenshots
Enock et al, 2014 [44]	Cognitive training via smartphone with attention bias modification training	An active control group and waitlist control	AA	Yes	No	Thorough descriptions, links, and few screenshots
Roepke et al, 2015 [45]	*SuperBetter:* Self-help game using either specific CBT or a general version using self-esteem and acceptance	2 versions of the app and 1 waitlist control	AA	Yes	3 authors work for SuperBetter (1 founded it)	Short description and 2 screenshots
Miner et al, 2016 [46]	*PTSD coach:* Psychoeducation, symptom assessments, self-management + access to supportive others	Waitlist control	AA	N/A	N/A	The app is available free online
Possemato et al, 2016 [47]	*PTSD coach:* Multifunctional psychoeducative self-help app with clinical support of 4, 20-min sessions	App alone	BT<=>AA	No	No	The app is available free online
Kuhn et al, 2017 [48]	*PTSD coach:* psychoeducation, symptom assessments, self-management + access to supportive others	Waitlist control	AA	No	N/A	The app is available free online

[a]TAU: Treatment as usual.

[b]Information assessed by author affiliation, grand support, and conflict of interest.

[c]CBT: cognitive behavioral therapy.

[d]N/A: not applicable.

[e]Two treatments not technologically attached but based on same therapy.

[f]Allowed to discuss the app with treating physician.

Table 4. Summarized findings from systematic review on 27 randomized controlled trials involving smartphones in the field of psychiatry. Consolidated Standards of Reporting Trials electronic health checklist is used as a guideline to systematically display trial design, methodology, and reporting of the identified trials.

Consort item	*Summarized findings according to CONSORT item, with references to relevant articles.*
Title and abstract (1a and 1b)	All but 6 titles [22,35,37,38,41,42] described the mode of delivery, components of treatment, target group, and trial design according to Consolidated Standards of Reporting Trials electronic health guidelines [20]. Often only broad terms of components were used, such as “mobile” or “mHealth.”
Introduction (2a and 2b)	The trials were published from 2013 to 2019 with equal distribution through 2013 to 2017 and increasing numbers from the from 2018 and forward [33-43]. Trials were mainly from western countries, especially from Scandinavia [23,25,27-30,34,40].
Trial design (3)	A total of 19 trials were classic RCTs [23,25-27,29-34,38-41,43-45,48], 7 were pilot RCTs [22,35-37,42,46,47], and 1 was a noninferiority RCT [28].
Participants (4a and 4b)	A total of 22 trials used research-based diagnoses: 5 were based on questionnaires [44-48], 8 used phone interviews [2225,28-30,34,42], 1 used Facetime or Skype interview [38], and 8 used personal interviews [26,27,31,32,36,39,43] mostly based on either MINI or DSM-4; 5 trials based their diagnoses only on clinical-based information [24,33,35,37,41]; 13 trials excluded patients with various degree of suicidal ideation [22,23,25,28-32,34,36,39,42,47]; 3 trials excluded patients with too severe symptomatology within the diagnosis of interest [22,26,27]. Most trials excluded patients with severe psychiatric comorbidity from lower International Statistical Classification of Diseases and Related Health Problems-10 chapter. In addition, 12 trials supplied participants with smartphones, either voluntary or mandatory [24,26,27,31,33,35,38,40,41,46-48]; 12 trials compensated participation in assessments with money or gift cards [26,33-38,42,43,46-48].
Interventions (5)	Intervention length varied substantially: from 3 weeks [36] to 52 weeks [41]. Most interventions lasted between 4 and 12 weeks; 8 trials used unaffected standard treatment beside intervention [24,27,35,37,38,40,44,45]; 8 trials tested the app alone [35-37,44-48], the remaining used variations of blended treatment; 1 trial tested blended therapy against app alone [47]; 18 trials used prompts to engage users, either from the app or by investigators [22,26-28,30,32-44]; 1 trial compared with an inactive “placebo” version of the app [41]; 2 trials compared with a placebo training module [43,44]. In 1 trial, participants in the control group received a “placebo” smartphone without the app system [27]; 5 trials used standard treatment as comparator [24,27,35,38,40], 11 trials used waitlist control [29,30,34,38,39,42-46,48], 4 trials used clinical intervention [26,28,31,33], and 5 trials compared with another app [23,25,35,36,42]; 1 trial collected automatically generated data [40] and further 8 trials collected data on app usage [24,33,34,42,43,45,47,48]; in 3 trials, intervention was only for iPhone [32,34,45] and in 3 only for Android [27,42,43]. The rest of the trials either had a Web-based version available or app for both platforms. Only 1 article mentioned information about updates of apps or intervention [27].
Outcomes (6a and 6b)	Overall, 8 trials did not use a predefined hierarchy of outcome measures [22,37,41-44,46,47]; 1 trial used tested levels of drug use in urine as a specific detection [37]; 1 trial used objective measurements of feasibility, use, and attrition as the primary outcome [35]; 1 trial tested a specific task [38]; 1 trial used video call–based clinical ratings [41]. Only 4 trials used clinical ratings as the primary outcome [26,27,31,40]. Remaining trials used patient-reported outcome measures. A total of 12 trials used internet platform for data collection [23,25,28-30,34,38,39,42-46,48], with 6 of these mentioning validations of questionnaires for online use [23,25,28,30,34,46]; 2 trials used the app for outcome measure [31,41].
Sample size (7a and 7b)	Sample size varied from 20 participants [47] to 429 participants [44]; 11 trials with numbers above 100 participants [24,30,32-34,39-41,44,45,48]. Pilot trials were smaller.
Randomization (8, 9, and 10)	Overall, 8 trials did not supply information about randomization [29,31,36-38,45-47]; 1 used Excel [43], and 1 used the app for randomization [41]. The remaining mainly used online software.
Blinding (11a and 11b)	Overall, 2 trials claim to be double-blinded with no further explanation on how blinding was assured [43,44]; 1 trial blinded app allocation for the patients (they tested 2 different apps) [45]. The remaining trials had no blinding of patients. In 12 trials, authors explicitly wrote that they used blinded assessments for outcome measures [25-29,31-33,35,38,40,44]. Within these 12 trials, 5 trials used patient-reported outcome measures as the primary outcome measure with nonblinded patients [25,28,29,32,33]; 1 trial tested for the success of blinding [32].
Statistical methods (12a and 12b)	A total of 11 trials based sample size on power calculations [24,27,28,30-33,39,40,45,48]. All but 1 of these managed to recruit at least the desired number [31]. No trials took changes and updates of software or technical problems into account in statistical methods.
Participant flow (13a and 13b)	All but 2 trials [43,44] presented a trial flow chart of eligible subjects, although with various details on reasons not to participate and drop out. Completion rates were reported very differently, varying from 163/164 completing primary outcome [32] to 74/283 completing posttreatment assessments (primary outcome) [45]. All but 2 trials reported on adherence to treatment [36,43].
Recruitment (14a and 14b)	Recruitment length was reported in 16 trials and varied from a few months to several years [22-24,27,30,32,33,35,37,41,42,44-48]; 10 trials used closed recruitment with a referral from clinicians or researchers seeking out participants from a well-defined patient population [24,27,31-33,35,37,40,41,47]; 1 trial gave no information on recruitment [43]. The remaining trials used open recruitment mainly via Craigslist.org, advertising in traditional ways, or through social media.
Baseline data (15)	Only 2 trials included technology-specific baseline data or information about participant technological abilities [32,33].
Numbers analyzed (16)	All but 6 trials [22,36,37,41-43] used the intent-to-treat principles in the primary analysis.

Consort item	*Summarized findings according to CONSORT item, with references to relevant articles.*
Outcome and estimation (17)	A total of 17 trials presented intensity of use or user data, either in the article or in supplementary data, with significant variations in usage among subjects and between trials [22-25,27,32-35,37-40,42,43,45,47].
Harms (19)	Overall, 5 trials prospectively measured harms or adverse events and reported directly in paper [32,33,35,41,42]. These trials found no harm from smartphone treatment used. One of these trials [42] had a safety protocol with clear, standardized instructions on how to react to suicidal ideation; 1 trial found a negative effect of treatment in secondary analysis, indicating fewer improvements in symptoms in a subgroup with a higher baseline score on the Hamilton Rating Scale compared with controls [27].No trials mentioned privacy breaches. Three trials mentioned technical problems and how these affected the intervention [23,40,48].
Generalizability (21)	Trials were heterogeneous. Some had strict criteria on diagnosis, comorbidity, and ongoing treatment, whereas others were pragmatic trials with few exclusion criteria. Trial populations varied from patients recruited among the general population who might not have sought help in the regular treatment system [22,23,25,28,30,34,39,42,44], whereas others came from specialized clinical functions setups [24,27,31-33,35,37,40,41,47].
Registration (23)	A total of 14 articles included information about trial registration [22,24-27,29-33,35,39-41].
Protocol (24)	A total of 5 trials published their trial protocol [27,30,32,35,40]; 1 trial had the protocol attached to the publication [28].
Funding (25)	Most authors came from universities; 15 trials reported information regarding funding [24-28,31,32,35-38,40,41,43,47] with funding mainly coming from public funds and institutions.
Competing interest (X27)[a]	A total of 9 articles declared having various degree of affiliation with private technology companies or closed relation to the app that they tested [25,30-34,40,42,45]; 8 trials did not include conflicts of interest in the printed article [23,28,29,35,37,39,46,48].

[a]Not an original Consolidated Standards of Reporting Trials item but included in the Consort electronic health checklist as *X27*.

Adherence to Smartphone Intervention

All but 2 trials [36,41] reported adherence by using one of the following or a combination: report on how many in-app lessons, modules or sessions participants completed, the number of daily or weekly users or logins, numbers of active users at a given time point, and composite scores measuring adherence.

Some trials included large tables on the usage of different components of the app, for example, the study by Schlosser et al [38], or included detailed program usage in the supplementary section, for example, the study by Stolz et al [39].

One trial economically compensated the participants for using the app [35]. Further comparison of adherence between trials was not feasible because of significant differences in how adherence was collected and reported in the individual trials.

Statistical Power Analysis

Overall, 11 trials based their required sample size on power analyses [24,27,28,30-33,39,40,45,48]. All but 1 [31] of these trials managed to recruit at least the desired number. In addition, 6 of the trials without power analyses were specified as pilot studies [22,35-37,42,46,47].

Statistical Analysis

All but 6 trials [22,36,37,41-43] used intent-to-treat principles in the primary analyses. Different methods were used to account for missing data. Some trials used multiple imputations [23,46,48], whereas others used mixed models [22,24,27,28,30,32,34,39,40,45] with *missing at random* or maximum likelihood estimations. One trial omitted data without imputation [41]. In addition, 8 trials included no explicit information on how the authors handled missing data [33,35-38,42,44,47]. Furthermore, 4 trials referred to an available predefined plan of analyses in the statistical section [27,32,33,35].

Technical Aspects of Smartphone-Based Treatment Interventions

One article informed about technical updates of the app, changes in interventions because of improvements, or technological problems that might have influenced the intervention over time [27]. None of the trials mentioned privacy breaches. In addition, 3 trials mentioned technical problems and how these affected interventions and the main hypothesis [23,40,48]. One trial collected automatically generated smartphone data (phone usage, social activity, and mobility) [40]. Furthermore, 8 trials collected data on app usage [24,33,34,42,43,45,47,48]. None of the trials accounted for changes or updates in the software or technical problems related to the intervention in the statistical analyses. Two trials included technology-specific baseline data or information about the participant's technological abilities such as prior smartphone use or assessments of the participant's ability to use a smartphone [32,33]. A total of 17 trials presented the intensity of use or user data with significant variations in usage among subjects and between trials [22-25,27,32-35,37-40,42,43,45,47].

Ethical Aspects

None of the included trials addressed the potential ethical aspects of using smartphones in the treatment of patients with a psychiatric disorder. Two trials included a section on ethics, including information on various trial registrations and approvals, data storage, and economic compensation [27,40].

Discussion

Principal Findings

This is the first systematic review regarding methodological challenges in RCTs investigating smartphone-based treatment interventions in patients with a psychiatric diagnosis. We included 27 trials with a wide range of psychiatric diagnoses and observed substantial between-trial heterogeneity. The trials

were conducted in diverse settings and used different smartphone-based treatment interventions and different follow-up periods. The trials reported on various outcome measures, which, in nearly half of the trials, were not clearly predefined. Most trials only used unblinded patient-evaluated outcome measures. A single trial reported on the success of blinding procedures.

Furthermore, only 1 trial provided information regarding technological updates of the smartphone-based treatment intervention. A declaration of interests was missing in 9 of the trials. No trial compared participants with nonparticipants, thereby increasing the risk of selection bias and making generalization of the trial findings difficult.

The included trials used smartphones in various ways and applied treatments to very heterogenic populations. Generally, the combination of insecure diagnoses, lack of blinding, use of patient-evaluated outcome measures, and lack of trial protocols or thorough publicly available trial registrations implies that evidence on the effects and side effects is still warranted.

The included trials used very different comparators, and the treatments given, besides the intervention of interest, varied from nothing to intensive clinical standard treatment setups. Generally, important aspects concerning technological features and how these inevitably affect outcomes were sparingly reported. On the basis of the results of the review, in the following sections, we discuss the highlights and suggest recommendations for designing and conducting future RCTs investigating the effect of smartphones in psychiatry.

Inclusion Criteria of Trial Patients

Owing to the lack of diagnostic biomarkers within psychiatry, currently, the research-based clinical diagnostic process represents the golden standard. The use of online, patient-evaluated diagnoses will reduce the validity of the diagnoses and reduce the generalizability to the clinical practice of trial results.

The use of a research-based clinical diagnostic assessment such as the Schedules for Clinical Assessment in Neuropsychiatry providing Diagnostic and Statistical Manual of Mental Disorders and ICD-10 diagnosis [49] or another systematic diagnostic assessment system should be prioritized. If patients are thoroughly and validly diagnosed and characterized before inclusion in the RCT, that is, by their treating doctors or psychologists, the clinical diagnoses may as well be used, depending on the aims and hypotheses of the RCT.

Smartphones allow for a vast amount of data to be collected online or automatically. This could be used as an advantage when conducting large RCTs. Digital solutions let researchers reach more participants and make it easier for patients to participate in clinical trials, not having to show up in the research settings for diagnostic assessment, baseline data collection, and outcome assessments. This possibly facilitates broader and more feasible inclusion of patients. Questionnaires conducted on the screen of a smartphone might vary compared with validated paper-based questionnaires. However, a Cochrane review from 2015 concludes that "apps might not affect data equivalence as long as the intended clinical application of the survey questionnaire, its intended frequency of administration and the setting in that it was validated remain unchanged" [50]. Telephonic interviews are different from clinical evaluations conducted by clinicians with possibly more reduced validity of diagnoses and sociodemographic and clinical data. Nevertheless, telephonic interviews (by trained lay interviewers or professionals) have been used in the research of depression and anxiety disorders with reasonable validity of diagnosis [51].

Modern electronic data collection should be used wisely with an awareness of possible changes in the quality of the data that the researchers obtain.

Interventions and Comparators

This systematic review showed that a placebo-controlled design was used to a limited extent. There is no clear definition of what a digital placebo treatment can or should contain, and defining a proper placebo group in a nonpharmacological RCTs is always difficult. The expectation for technical solutions themselves to be helpful resembles a reaction that is comparable to receiving a placebo pill and has been suggested to affect patients independent of active treatment [52]. The term *digital placebo* has been suggested [52]. Still, little is known about this issue; however, ongoing RCTs are investigating this subject further using a sham app as a comparator [53].

Clear descriptions of the content of the interventions and comparators used by researchers in future RCTs should be prioritized and made available to readers. If not mentioned in the primary publication, clear reference to the description should be made, such as in the study protocol, earlier publication, appendices, or publicly available versions of the intervention used. Furthermore, researchers could beneficially design the active intervention to fit into clinical practice and adapt to the clinical settings either as a stand-alone treatment or in combination with clinical treatment and support.

Outcome Measures and Power Analyses

A clear predefined research question, represented by a predefined primary outcome measure in a precisely well-defined patient population, is necessary. Selecting a predefined and relevant primary outcome measure is crucial [12]. Assessors should be blinded to outcome measures. End points of clinical relevance, that is independent of or blinded to researchers, such as admittance to a hospital or relapse or recurrence of illness (as in an ongoing trial [54]), should be prioritized compared with biased end points [12]. Such outcomes benefit from being critical for patients, relatives, and clinicians.

Power analyses should preferably be made before start the trial to ensure that the required sample is realistic and able to answer the primary research question—often represented by significant changes in the primary outcome.

When using clinician-based outcome measures, this should preferably be done by blinded trained clinical researchers. This is a difficult task because patients may give the researcher a hint on their allocation. Precautions to hinder this should be taken and described.

Although patient-evaluated outcome measures are appealing because of the ease use, we should be careful when interpreting

findings based on unblinded patient-evaluated outcomes as there will be a risk of bias.

Clinicians and administrators may play important roles in implementing such treatment systems following trial findings if the results indicate a possible effect of the new treatment [55].

Possible outcomes could include measures of clinicians' and caregivers' attitudes toward the intervention if involved. Further data regarding the use of resources and economic costs related to the intervention could be included if available.

Publication of a Trial Protocol and Reporting Guidelines

Compilation and publication of a trial protocol or a thorough description on publicly available registration sites, such as https://clinicaltrials.gov/, before analyzing data from the trial will increase the validity of the findings. Trial protocols and registrations should follow the standards for medical trial protocols [20] such as precise descriptions of inclusion and exclusion criteria, recruitment procedures, prioritized outcome measures including unblinded assessment, statistical power analyses, and a plan for analyses of participants versus nonparticipants. Furthermore, a thorough description of the used technology should be included [54,56,57]. Authors should publish deviations on especially the technical side of the protocol that were necessary, including valid arguments for the changes.

When reporting results, guidelines such as the CONSORT eHealth checklist [20] should be followed to increase transparency and for evaluation of findings.

Multidisciplinary Research

A close collaboration between information technology developers, clinicians and researchers is critical to ensure that the technology developed can answer the primary questions that are addressed and, vice versa, that the posed hypotheses match the available technology.

Authors should use the same transparency when working with the technology industry as researchers, doctors, and funders use with the medical industry. This systematic review shows that the authors' declarations of conflicts of interest are frequently missing. The industry has a significant direct or indirect interest in developing app solutions and gathers information from patients as information in the digital economy involves major economic interests [58]. Any potential conflicts of interest should always be declared.

Technology in Randomized Controlled Trials

Smartphones represent a unique tool to measure adherence and fidelity in RCTs. However, there is no clear definition of how to measure or report adherence to smartphone-based treatment systems.

Smartphone solutions might assist in follow-up assessments and could advantageously use built-in features to ease the patients. Even when patients are "lost to follow-up," assessments based on smartphone sensors could be collected, with patient permission, and might reflect behavioral changes for the user [59-61].

Prompts and reminders may comprise an important part of digital solutions to keep patients engaged in treatment.

These prompts, preferably, should be a part of the intended intervention that is investigated and not done only as part of the research project. The research team will potentially not be part of a real-world implementation of the technology and therefore give a false picture of adherence and effect of the intervention.

Information about technical skills and smartphone usage and habits might be valuable information and should be presented.

Reporting possible harms and adverse events are essential as the field of smartphone-based treatment is new, and little is known of potentially harmful effects such as worsening of symptoms or suicidal ideation. One trial has reported possible harms of using the intervention [27].

Potential harms and safety matters should be taken into consideration when designing and conducting the trial, as the example of the study by Stiles-Shields et al, [42] in which a safety protocol was used to standardize how researchers should react on suicidal ideations, expressed by participants either in the app or as part of the online questionnaires (where suicidal thoughts is a common item).

Updates, Revisions, and Adaptive Trials

Developing well-designed RCTs, including proper diagnostic procedures and thorough, robust, and blinded outcome measures, will inevitably be a time-consuming process. Still, it is crucial to test new smartphone-based treatment interventions on well-defined patient populations. *Locking* a digital intervention for several years is a challenge because technology would be expected to evolve and improve over time. In software engineering, updates and enhancements of the software are frequently released to fix bugs and errors and to improve usability, stability, robustness, and security. Moreover, the type of data collected from the phone via sensors and its processing by algorithms will also constantly change and adapt according to automatic machine learning models. Such learning systems will change based on new information and might be used increasingly in psychiatry as well as in many other fields [62]. Finally, the hardware and operating systems of the mobile phone regularly change, and the major vendors (Apple and Google) regularly release new hardware and operating systems that affect what an app can and cannot do.

Consequently, the concept of maintaining the digital intervention stable during a clinical trial is unrealistic, taking into consideration the long time span involved in conducting most RCTs.

To merge these opposing interests, we must accept these ongoing changes as a natural consequence of conducting research in the field, similar to surgeons who improve their skills or therapist who adapt and improve as they see more and more patients. Such updates should be reported systematically and should be thoroughly described so that it will be clear to readers which type of updates were made and what effect they might have had on trial results. In addition, the possible effects of add-ons could be presented and provided enough statistical power for subanalyses.

By allowing a more flexible approach to the RCT design, it would conform more to the so-called adaptive trial paradigm [17,63], and then RCTs might be able to answer questions regarding the effect of mHealth interventions without evaluating outdated technology [16,17].

It is necessary to accept and incorporate technological developments and changes in the design of a specific RCT in line with suggestions from similar but older scientific fields as telehealth in general [64].

Ethical Aspects

In a digital economy, data generated by apps and services constitute a valuable resource. Within psychiatry and other areas, such data can lead to discrimination of individuals and patient groups as a whole, and therefore, the handling and use of the data collected are of great importance. The value such information has to industry and other stakeholders is driving the many apps that are widely available without evidence of effects or possible harms. We need information obtained through well-designed and transparent trials to improve our knowledge of how these treatments can help patients and professionals.

Privacy, security, technology illiteracy, depersonalization of treatment systems, and technological paternalism are some of many possible ethical issues in this field [58].

These challenges apply both to the trial designs and the treatments being developed. The included trials hardly mentioned any such ethical issues.

Trials must consider and discuss possible ethical implications of the trial design and of the treatment itself.

Findings in a Scientific Context

The results from one of the included trials have been replicated using the same intervention, trial design, and outcome measures in a second trial only once [27,40]. However, this is necessary to increase evidence. Notably, numerous trial protocols within the field were identified during the search [54,65-69]. This indicates that the pace of the field is still increasing, and hopefully, future trials will provide more evidence within the field of smartphone-based treatment in patients with psychiatric disorders. This development is further shown in the large numbers of RCTs found in this review. Only 1 of the included trials in this review was published and included in a previous review from 2013 investigating the effects of smartphones in mental health [18]. With an increasing number of RCTs and hopefully increasing the quality of design and method, and a more uniform reporting of results, future meta-analyses of the effects of treatments will be possible.

Limitations

This study has various limitations: not all parts of the review process and data collection were double checked, no protocol of study method was published online beforehand, and no meta-analyses or statistical analyses were performed on the included data. We did not use any scoring systems for the assessments of quality and risk of bias in the included studies, leaving the assessment to the reader.

Conclusions

This first systematic review on the design, conduct, and methodological challenges of RCTs investigating the effect of smartphone-based treatment in patients with a psychiatric diagnosis suggests that there is a rapidly increasing interest for this type of treatment. Although an increasing number of trials tested new smartphone-based treatments, the trial designs and reporting were of low quality compared with more classic medical RCTs, and heterogeneity and methodological issues in individual trials limit the evidence.

Smartphone-based treatment interventions imply new challenges and opportunities, but we, as researchers, should consider strict methodological efforts when designing, conducting, and reporting such trials as in the rest of the field of medicine. Future trials employing strict methodology, including detailed description regarding patient recruitment, pre- and well-defined, prioritized outcome measures, information regarding technical updates and down periods, and statements on potential conflicts of interest are warranted. Research groups without trial experience should seek out information on how to conduct RCTs with high a methodological standard to ensure a high level of quality in the research. Finally, close collaborations between professions and specialties are needed in this complex branch of science.

Acknowledgments

MLT's work is supported by the Innovation Foundation (5164-00001B).

Conflicts of Interest

All authors are currently working on the RADMIS project, which is the successor of the MONARCA system. LVK has been a consultant for Sunovion in the past 3 years. JEB is a cofounder and chief scientific officer in MONSENSO, a technology company developing and selling app-based solutions for psychiatric diseases. The other authors have no conflicts to declare.

Multimedia Appendix 1
Data extraction template.
[PDF File (Adobe PDF File), 116 KB - jmir_v21i10e15362_app1.pdf]

References

1. GBD 2017 Disease and Injury Incidence and Prevalence Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet 2018 Nov 10;392(10159):1789-1858 [FREE Full text] [doi: 10.1016/S0140-6736(18)32279-7] [Medline: 30496104]
2. Trautmann S, Rehm J, Wittchen H. The economic costs of mental disorders: do our societies react appropriately to the burden of mental disorders? EMBO Rep 2016 Sep;17(9):1245-1249 [FREE Full text] [doi: 10.15252/embr.201642951] [Medline: 27491723]
3. Wittchen H, Jacobi F. Size and burden of mental disorders in Europe--a critical review and appraisal of 27 studies. Eur Neuropsychopharmacol 2005 Aug;15(4):357-376. [doi: 10.1016/j.euroneuro.2005.04.012] [Medline: 15961293]
4. World Health Organization. Fact Sheet on Mental Disorders URL: https://www.who.int/campaigns/world-health-day/2017/fact-sheets/en/ [accessed 2018-12-13]
5. World Health Organization. mHealth: New Horizons for Health Through Mobile Technologies. Geneva, Switzerland: World Health Organization; 2011.
6. Statista. Number of Smartphone Users Worldwide From 2016 to 2021 (in Billions) URL: https://www.statista.com/statistics/330695/number-of-smartphone-users-worldwide/ [accessed 2018-12-13]
7. Netimperative. 2017. Denmark 'Has Highest Smartphone Penetration Rate in the World' URL: http://www.netimperative.com/2017/12/denmark-highest-smartphone-penetration-rate-world/ [accessed 2018-12-13]
8. Dogan E, Sander C, Wagner X, Hegerl U, Kohls E. Smartphone-based monitoring of objective and subjective data in affective disorders: where are we and where are we going? Systematic review. J Med Internet Res 2017 Jul 24;19(7):e262 [FREE Full text] [doi: 10.2196/jmir.7006] [Medline: 28739561]
9. Faurholt-Jepsen M, Munkholm K, Frost M, Bardram JE, Kessing LV. Electronic self-monitoring of mood using IT platforms in adult patients with bipolar disorder: a systematic review of the validity and evidence. BMC Psychiatry 2016 Jan 15;16:7 [FREE Full text] [doi: 10.1186/s12888-016-0713-0] [Medline: 26769120]
10. Wang K, Varma DS, Prosperi M. A systematic review of the effectiveness of mobile apps for monitoring and management of mental health symptoms or disorders. J Psychiatr Res 2018 Dec;107:73-78. [doi: 10.1016/j.jpsychires.2018.10.006] [Medline: 30347316]
11. Anthes E. Mental health: there's an app for that. Nature 2016 Apr 7;532(7597):20-23. [doi: 10.1038/532020a] [Medline: 27078548]
12. Schulz KF, Grimes DA. Generation of allocation sequences in randomised trials: chance, not choice. Lancet 2002 Feb 9;359(9305):515-519. [doi: 10.1016/S0140-6736(02)07683-3] [Medline: 11853818]
13. Strand LB, Clarke P, Graves N, Barnett AG. Time to publication for publicly funded clinical trials in Australia: an observational study. BMJ Open 2017 Mar 22;7(3):e012212 [FREE Full text] [doi: 10.1136/bmjopen-2016-012212] [Medline: 28336734]
14. Riley WT, Glasgow RE, Etheredge L, Abernethy AP. Rapid, responsive, relevant (R3) research: a call for a rapid learning health research enterprise. Clin Transl Med 2013 May 10;2(1):10 [FREE Full text] [doi: 10.1186/2001-1326-2-10] [Medline: 23663660]
15. Mohr DC, Schueller SM, Riley WT, Brown CH, Cuijpers P, Duan N, et al. Trials of intervention principles: evaluation methods for evolving behavioral intervention technologies. J Med Internet Res 2015 Jul 8;17(7):e166 [FREE Full text] [doi: 10.2196/jmir.4391] [Medline: 26155878]
16. Baker TB, Gustafson DH, Shah D. How can research keep up with eHealth? Ten strategies for increasing the timeliness and usefulness of eHealth research. J Med Internet Res 2014 Feb 19;16(2):e36 [FREE Full text] [doi: 10.2196/jmir.2925] [Medline: 24554442]
17. Law LM, Wason JM. Design of telehealth trials--introducing adaptive approaches. Int J Med Inform 2014 Dec;83(12):870-880 [FREE Full text] [doi: 10.1016/j.ijmedinf.2014.09.002] [Medline: 25293533]
18. Donker T, Petrie K, Proudfoot J, Clarke J, Birch M, Christensen H. Smartphones for smarter delivery of mental health programs: a systematic review. J Med Internet Res 2013 Nov 15;15(11):e247 [FREE Full text] [doi: 10.2196/jmir.2791] [Medline: 24240579]
19. Firth J, Torous J, Nicholas J, Carney R, Pratap A, Rosenbaum S, et al. The efficacy of smartphone-based mental health interventions for depressive symptoms: a meta-analysis of randomized controlled trials. World Psychiatry 2017 Oct;16(3):287-298 [FREE Full text] [doi: 10.1002/wps.20472] [Medline: 28941113]
20. Eysenbach G, CONSORT-EHEALTH Group. CONSORT-EHEALTH: improving and standardizing evaluation reports of web-based and mobile health interventions. J Med Internet Res 2011 Dec 31;13(4):e126 [FREE Full text] [doi: 10.2196/jmir.1923] [Medline: 22209829]
21. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Br Med J 2009 Jul 21;339:b2535 [FREE Full text] [doi: 10.1136/bmj.b2535] [Medline: 19622551]
22. Watts S, Mackenzie A, Thomas C, Griskaitis A, Mewton L, Williams A, et al. CBT for depression: a pilot RCT comparing mobile phone vs computer. BMC Psychiatry 2013 Feb 7;13:49 [FREE Full text] [doi: 10.1186/1471-244X-13-49] [Medline: 23391304]

23. Dagöö J, Asplund RP, Bsenko HA, Hjerling S, Holmberg A, Westh S, et al. Cognitive behavior therapy versus interpersonal psychotherapy for social anxiety disorder delivered via smartphone and computer: a randomized controlled trial. J Anxiety Disord 2014 May;28(4):410-417. [doi: 10.1016/j.janxdis.2014.02.003] [Medline: 24731441]
24. Gustafson DH, McTavish FM, Chih M, Atwood AK, Johnson RA, Boyle MG, et al. A smartphone application to support recovery from alcoholism: a randomized clinical trial. JAMA Psychiatry 2014 May;71(5):566-572 [FREE Full text] [doi: 10.1001/jamapsychiatry.2013.4642] [Medline: 24671165]
25. Ly KH, Trüschel A, Jarl L, Magnusson S, Windahl T, Johansson R, et al. Behavioural activation versus mindfulness-based guided self-help treatment administered through a smartphone application: a randomised controlled trial. BMJ Open 2014 Jan 9;4(1):e003440 [FREE Full text] [doi: 10.1136/bmjopen-2013-003440] [Medline: 24413342]
26. Depp CA, Ceglowski J, Wang VC, Yaghouti F, Mausbach BT, Thompson WK, et al. Augmenting psychoeducation with a mobile intervention for bipolar disorder: a randomized controlled trial. J Affect Disord 2015 Mar 15;174:23-30 [FREE Full text] [doi: 10.1016/j.jad.2014.10.053] [Medline: 25479050]
27. Faurholt-Jepsen M, Frost M, Ritz C, Christensen EM, Jacoby AS, Mikkelsen RL, et al. Daily electronic self-monitoring in bipolar disorder using smartphones - the MONARCA I trial: a randomized, placebo-controlled, single-blind, parallel group trial. Psychol Med 2015 Oct;45(13):2691-2704. [doi: 10.1017/S0033291715000410] [Medline: 26220802]
28. Ly KH, Topooco N, Cederlund H, Wallin A, Bergström J, Molander O, et al. Smartphone-supported versus full behavioural activation for depression: a randomised controlled trial. PLoS One 2015;10(5):e0126559 [FREE Full text] [doi: 10.1371/journal.pone.0126559] [Medline: 26010890]
29. Moëll B, Kollberg L, Nasri B, Lindefors N, Kaldo V. Living SMART — a randomized controlled trial of a guided online course teaching adults with ADHD or sub-clinical ADHD to use smartphones to structure their everyday life. Internet Interv 2015 Mar;2(1):24-31. [doi: 10.1016/j.invent.2014.11.004]
30. Ivanova E, Lindner P, Ly KH, Dahlin M, Vernmark K, Andersson G, et al. Guided and unguided acceptance and commitment therapy for social anxiety disorder and/or panic disorder provided via the internet and a smartphone application: a randomized controlled trial. J Anxiety Disord 2016 Dec;44:27-35. [doi: 10.1016/j.janxdis.2016.09.012] [Medline: 27721123]
31. Hildebrandt T, Michaelides A, Mackinnon D, Greif R, DeBar L, Sysko R. Randomized controlled trial comparing smartphone assisted versus traditional guided self-help for adults with binge eating. Int J Eat Disord 2017 Nov;50(11):1313-1322 [FREE Full text] [doi: 10.1002/eat.22781] [Medline: 28960384]
32. Mantani A, Kato T, Furukawa TA, Horikoshi M, Imai H, Hiroe T, et al. Smartphone cognitive behavioral therapy as an adjunct to pharmacotherapy for refractory depression: randomized controlled trial. J Med Internet Res 2017 Nov 3;19(11):e373 [FREE Full text] [doi: 10.2196/jmir.8602] [Medline: 29101095]
33. Ben-Zeev D, Brian RM, Jonathan G, Razzano L, Pashka N, Carpenter-Song E, et al. Mobile Health (mHealth) versus clinic-based group intervention for people with serious mental illness: a randomized controlled trial. Psychiatr Serv 2018 Sep 1;69(9):978-985. [doi: 10.1176/appi.ps.201800063] [Medline: 29793397]
34. Boettcher J, Magnusson K, Marklund A, Berglund E, Blomdahl R, Braun U, et al. Adding a smartphone app to internet-based self-help for social anxiety: a randomized controlled trial. Comput Hum Behav 2018 Oct;87(6):98-108. [doi: 10.1016/j.chb.2018.04.052]
35. Bucci S, Barrowclough C, Ainsworth J, Machin M, Morris R, Berry K, et al. Actissist: proof-of-concept trial of a theory-driven digital intervention for psychosis. Schizophr Bull 2018 Aug 20;44(5):1070-1080 [FREE Full text] [doi: 10.1093/schbul/sby032] [Medline: 29566206]
36. Hur J, Kim B, Park D, Choi S. A scenario-based cognitive behavioral therapy mobile app to reduce dysfunctional beliefs in individuals with depression: a randomized controlled trial. Telemed J E Health 2018 Sep;24(9):710-716. [doi: 10.1089/tmj.2017.0214] [Medline: 29323626]
37. Liang D, Han H, Du J, Zhao M, Hser Y. A pilot study of a smartphone application supporting recovery from drug addiction. J Subst Abuse Treat 2018 May;88:51-58 [FREE Full text] [doi: 10.1016/j.jsat.2018.02.006] [Medline: 29606226]
38. Schlosser DA, Campellone TR, Truong B, Etter K, Vergani S, Komaiko K, et al. Efficacy of PRIME, a mobile app intervention designed to improve motivation in young people with schizophrenia. Schizophr Bull 2018 Aug 20;44(5):1010-1020 [FREE Full text] [doi: 10.1093/schbul/sby078] [Medline: 29939367]
39. Stolz T, Schulz A, Krieger T, Vincent A, Urech A, Moser C, et al. A mobile app for social anxiety disorder: a three-arm randomized controlled trial comparing mobile and PC-based guided self-help interventions. J Consult Clin Psychol 2018 Jun;86(6):493-504. [doi: 10.1037/ccp0000301] [Medline: 29781648]
40. Faurholt-Jepsen M, Frost M, Christensen EM, Bardram JE, Vinberg M, Kessing LV. The effect of smartphone-based monitoring on illness activity in bipolar disorder: the MONARCA II randomized controlled single-blinded trial. Psychol Med 2019 Apr 4:- (epub ahead of print). [doi: 10.1017/S0033291719000710] [Medline: 30944054]
41. Krzystanek M, Borkowski M, Skałacka K, Krysta K. A telemedicine platform to improve clinical parameters in paranoid schizophrenia patients: results of a one-year randomized study. Schizophr Res 2019 Feb;204:389-396. [doi: 10.1016/j.schres.2018.08.016] [Medline: 30154027]
42. Stiles-Shields C, Montague E, Kwasny MJ, Mohr DC. Behavioral and cognitive intervention strategies delivered via coached apps for depression: pilot trial. Psychol Serv 2019 May;16(2):233-238. [doi: 10.1037/ser0000261] [Medline: 30407055]

43. Teng M, Hou Y, Chang S, Cheng H. Home-delivered attention bias modification training via smartphone to improve attention control in sub-clinical generalized anxiety disorder: a randomized, controlled multi-session experiment. J Affect Disord 2019 Mar 1;246:444-451. [doi: 10.1016/j.jad.2018.12.118] [Medline: 30599367]
44. Enock PM, Hofmann SG, McNally RJ. Attention bias modification training via smartphone to reduce social anxiety: a randomized, controlled multi-session experiment. Cogn Ther Res 2014 Mar 4;38(2):200-216. [doi: 10.1007/s10608-014-9606-z]
45. Roepke AM, Jaffee SR, Riffle OM, McGonigal J, Broome R, Maxwell B. Randomized controlled trial of SuperBetter, a smartphone-based/internet-based self-help tool to reduce depressive symptoms. Games Health J 2015 Jun;4(3):235-246. [doi: 10.1089/g4h.2014.0046] [Medline: 26182069]
46. Miner A, Kuhn E, Hoffman JE, Owen JE, Ruzek JI, Taylor CB. Feasibility, acceptability, and potential efficacy of the PTSD Coach app: a pilot randomized controlled trial with community trauma survivors. Psychol Trauma 2016 May;8(3):384-392. [doi: 10.1037/tra0000092] [Medline: 27046668]
47. Possemato K, Kuhn E, Johnson E, Hoffman JE, Owen JE, Kanuri N, et al. Using PTSD coach in primary care with and without clinician support: a pilot randomized controlled trial. Gen Hosp Psychiatry 2016;38:94-98. [doi: 10.1016/j.genhosppsych.2015.09.005] [Medline: 26589765]
48. Kuhn E, Kanuri N, Hoffman JE, Garvert DW, Ruzek JI, Taylor CB. A randomized controlled trial of a smartphone app for posttraumatic stress disorder symptoms. J Consult Clin Psychol 2017 Mar;85(3):267-273. [doi: 10.1037/ccp0000163] [Medline: 28221061]
49. Wing JK, Babor T, Brugha T, Burke J, Cooper JE, Giel R, et al. SCAN. Schedules for clinical assessment in neuropsychiatry. Arch Gen Psychiatry 1990 Jun;47(6):589-593. [doi: 10.1001/archpsyc.1990.01810180089012] [Medline: 2190539]
50. Belisario JS, Jamsek J, Huckvale K, O'Donoghue J, Morrison CP, Car J. Comparison of self-administered survey questionnaire responses collected using mobile apps versus other methods. Cochrane Database Syst Rev 2015 Jul 27(7):MR000042. [doi: 10.1002/14651858.MR000042.pub2] [Medline: 26212714]
51. Muskens EM, Lucassen P, Groenleer W, van Weel C, Oude Voshaar R, Speckens A. Psychiatric diagnosis by telephone: is it an opportunity? Soc Psychiatry Psychiatr Epidemiol 2014 Oct;49(10):1677-1689. [doi: 10.1007/s00127-014-0861-9] [Medline: 24632782]
52. Torous J, Firth J. The digital placebo effect: mobile mental health meets clinical psychiatry. Lancet Psychiatry 2016 Feb;3(2):100-102. [doi: 10.1016/S2215-0366(15)00565-9] [Medline: 26851322]
53. Giosan C, Cobeanu O, Mogoaşe C, Szentagotai A, Mureşan V, Boian R. Reducing depressive symptomatology with a smartphone app: study protocol for a randomized, placebo-controlled trial. Trials 2017 May 12;18(1):215 [FREE Full text] [doi: 10.1186/s13063-017-1960-1] [Medline: 28494802]
54. Faurholt-Jepsen M, Frost M, Martiny K, Tuxen N, Rosenberg N, Busk J, et al. Reducing the rate and duration of re-admissions among patients with unipolar disorder and bipolar disorder using smartphone-based monitoring and treatment - the RADMIS trials: study protocol for two randomized controlled trials. Trials 2017 Jun 15;18(1):277 [FREE Full text] [doi: 10.1186/s13063-017-2015-3] [Medline: 28619114]
55. Kerst A, Zielasek J, Gaebel W. Smartphone applications for depression: a systematic literature review and a survey of health care professionals' attitudes towards their use in clinical practice. Eur Arch Psychiatry Clin Neurosci 2019 Jan 03:- (epub ahead of print). [doi: 10.1007/s00406-018-0974-3] [Medline: 30607530]
56. Faurholt-Jepsen M, Vinberg M, Christensen EM, Frost M, Bardram J, Kessing LV. Daily electronic self-monitoring of subjective and objective symptoms in bipolar disorder--the MONARCA trial protocol (MONitoring, treAtment and pRediCtion of bipolAr disorder episodes): a randomised controlled single-blind trial. BMJ Open 2013;3(7):- [FREE Full text] [doi: 10.1136/bmjopen-2013-003353] [Medline: 23883891]
57. Faurholt-Jepsen M, Vinberg M, Frost M, Christensen EM, Bardram J, Kessing LV. Daily electronic monitoring of subjective and objective measures of illness activity in bipolar disorder using smartphones--the MONARCA II trial protocol: a randomized controlled single-blind parallel-group trial. BMC Psychiatry 2014 Nov 25;14:309 [FREE Full text] [doi: 10.1186/s12888-014-0309-5] [Medline: 25420431]
58. Bauer M, Glenn T, Monteith S, Bauer R, Whybrow PC, Geddes J. Ethical perspectives on recommending digital technology for patients with mental illness. Int J Bipolar Disord 2017 Dec;5(1):6 [FREE Full text] [doi: 10.1186/s40345-017-0073-9] [Medline: 28155206]
59. Faurholt-Jepsen M, Bauer M, Kessing LV. Smartphone-based objective monitoring in bipolar disorder: status and considerations. Int J Bipolar Disord 2018 Jan 23;6(1):6 [FREE Full text] [doi: 10.1186/s40345-017-0110-8] [Medline: 29359252]
60. Faurholt-Jepsen M, Frost M, Vinberg M, Christensen EM, Bardram JE, Kessing LV. Smartphone data as objective measures of bipolar disorder symptoms. Psychiatry Res 2014 Jun 30;217(1-2):124-127. [doi: 10.1016/j.psychres.2014.03.009] [Medline: 24679993]
61. Faurholt-Jepsen M, Brage S, Vinberg M, Jensen HM, Christensen EM, Knorr U, et al. Electronic monitoring of psychomotor activity as a supplementary objective measure of depression severity. Nord J Psychiatry 2015 Feb;69(2):118-125. [doi: 10.3109/08039488.2014.936501] [Medline: 25131795]

62. Bzdok D, Meyer-Lindenberg A. Machine learning for precision psychiatry: opportunities and challenges. Biol Psychiatry Cogn Neurosci Neuroimaging 2018 Mar;3(3):223-230. [doi: 10.1016/j.bpsc.2017.11.007] [Medline: 29486863]
63. Park JJ, Thorlund K, Mills EJ. Critical concepts in adaptive clinical trials. Clin Epidemiol 2018;10:343-351 [FREE Full text] [doi: 10.2147/CLEP.S156708] [Medline: 29606891]
64. Ekeland AG, Bowes A, Flottorp S. Methodologies for assessing telemedicine: a systematic review of reviews. Int J Med Inform 2012 Jan;81(1):1-11. [doi: 10.1016/j.ijmedinf.2011.10.009] [Medline: 22104370]
65. Mühlbauer E, Bauer M, Ebner-Priemer U, Ritter P, Hill H, Beier F, et al. Effectiveness of smartphone-based ambulatory assessment (SBAA-BD) including a predicting system for upcoming episodes in the long-term treatment of patients with bipolar disorders: study protocol for a randomized controlled single-blind trial. BMC Psychiatry 2018 Oct 26;18(1):349 [FREE Full text] [doi: 10.1186/s12888-018-1929-y] [Medline: 30367608]
66. Anastasiadou D, Lupiañez-Villanueva F, Faulí C, Arcal Cunillera J, Serrano-Troncoso E. Cost-effectiveness of the mobile application TCApp combined with face-to-face CBT treatment compared to face-to-face CBT treatment alone for patients with an eating disorder: study protocol of a multi-centre randomised controlled trial. BMC Psychiatry 2018 May 2;18(1):118 [FREE Full text] [doi: 10.1186/s12888-018-1664-4] [Medline: 29716580]
67. Scott CK, Dennis ML, Gustafson DH. Using smartphones to decrease substance use via self-monitoring and recovery support: study protocol for a randomized control trial. Trials 2017 Aug 10;18(1):374 [FREE Full text] [doi: 10.1186/s13063-017-2096-z] [Medline: 28797307]
68. Kolar DR, Hammerle F, Jenetzky E, Huss M. Smartphone-enhanced low-threshold intervention for adolescents with anorexia nervosa (SELTIAN) waiting for outpatient psychotherapy: study protocol of a randomised controlled trial. BMJ Open 2017 Oct 22;7(10):e018049 [FREE Full text] [doi: 10.1136/bmjopen-2017-018049] [Medline: 29061627]
69. Kemmeren LL, van Schaik DJF, Riper H, Kleiboer AM, Bosmans JE, Smit JH. Effectiveness of blended depression treatment for adults in specialised mental healthcare: study protocol for a randomised controlled trial. BMC Psychiatry 2016 Apr 21;16:113 [FREE Full text] [doi: 10.1186/s12888-016-0818-5] [Medline: 27102812]

Abbreviations

BDI-II: Beck Depression Inventory
CBT: cognitive behavioral therapy
CES-D: Center for Epidemiologic Studies Depression Scale
CONSORT: Consolidated Standards of Reporting Trials
DSM: Diagnostic and Statistical Manual of Mental Disorders
eHealth: electronic health
ICD-10: International Statistical Classification of Diseases and Related Health Problems-10
K-10: Kessler Psychological Distress Scale
LSAS-SR: Liebowitz Social Anxiety Scale—self-reported
MADRS: Montgomery–Åsberg Depression Rating Scale
mHealth: mobile health
MINI: Mini International Neuropsychiatric Interview
NI: noninferiority
PCL-C: Post-Traumatic Checklist—Civilian
PCL-S: Post-Traumatic Checklist Scale
PHQ-9: Patient Health Questionnaire-9
PTSD: posttraumatic stress disorder
RCT: randomized controlled trial
SCAN: Schedules for Clinical Assessment in Neuropsychiatry
SCID: Structured Clinical Interview.
SIAS: Social Interaction Anxiety Scale
TAU: Treatment as usual
YMRS: Young Mania Rating Scale

Edited by G Eysenbach; submitted 04.07.19; peer-reviewed by C Simons, D Ben-Zeev; comments to author 23.07.19; revised version received 28.08.19; accepted 04.09.19; published 27.10.19.

Please cite as:
Tønning ML, Kessing LV, Bardram JE, Faurholt-Jepsen M
Methodological Challenges in Randomized Controlled Trials on Smartphone-Based Treatment in Psychiatry: Systematic Review
J Med Internet Res 2019;21(10):e15362
URL: http://www.jmir.org/2019/10/e15362/
doi:10.2196/15362
PMID:31663859

Original Paper

Trustworthy Health-Related Tweets on Social Media in Saudi Arabia: Tweet Metadata Analysis

Yahya Albalawi[1,2,3], MSc; Nikola S Nikolov[1,3], PhD; Jim Buckley[1,3], PhD

[1]Department of Computer Science and Information Systems, University of Limerick, Limerick, Ireland

[2]Department of Computer and Information Sciences, College of Arts and Science, University of Taibah, Al-Ula, Saudi Arabia

[3]The Irish Software Research Centre, Lero, University of Limerick, Limerick, Ireland

Corresponding Author:
Yahya Albalawi, MSc
Department of Computer Science and Information Systems
University of Limerick
Tierney Building
Limerick, V94 NYD3
Ireland
Phone: 353 61213028 ext 3724
Email: yahalbalawi@gmail.com

Abstract

Background: Social media platforms play a vital role in the dissemination of health information. However, evidence suggests that a high proportion of Twitter posts (ie, tweets) are not necessarily accurate, and many studies suggest that tweets do not need to be accurate, or at least evidence based, to receive traction. This is a dangerous combination in the sphere of health information.

Objective: The first objective of this study is to examine health-related tweets originating from Saudi Arabia in terms of their accuracy. The second objective is to find factors that relate to the accuracy and dissemination of these tweets, thereby enabling the identification of ways to enhance the dissemination of accurate tweets. The initial findings from this study and methodological improvements will then be employed in a larger-scale study that will address these issues in more detail.

Methods: A health lexicon was used to extract health-related tweets using the Twitter application programming interface and the results were further filtered manually. A total of 300 tweets were each labeled by two medical doctors; the doctors agreed that 109 tweets were either accurate or inaccurate. Other measures were taken from these tweets' metadata to see if there was any relationship between the measures and either the accuracy or the dissemination of the tweets. The entire range of this metadata was analyzed using Python, version 3.6.5 (Python Software Foundation), to answer the research questions posed.

Results: A total of 34 out of 109 tweets (31.2%) in the dataset used in this study were classified as untrustworthy health information. These came mainly from users with a non-health care background and social media accounts that had no corresponding physical (ie, organization) manifestation. Unsurprisingly, we found that traditionally trusted health sources were more likely to tweet accurate health information than other users. Likewise, these provisional results suggest that tweets posted in the morning are more trustworthy than tweets posted at night, possibly corresponding to official and casual posts, respectively. Our results also suggest that the crowd was quite good at identifying trustworthy information sources, as evidenced by the number of times a tweet's author was tagged as *favorited* by the community.

Conclusions: The results indicate some initially surprising factors that might correlate with the accuracy of tweets and their dissemination. For example, the time a tweet was posted correlated with its accuracy, which may reflect a difference between professional (ie, morning) and hobbyist (ie, evening) tweets. More surprisingly, tweets containing a kashida—a decorative element in Arabic writing used to justify the text within lines—were more likely to be disseminated through retweets. These findings will be further assessed using data analysis techniques on a much larger dataset in future work.

(J Med Internet Res 2019;21(10):e14731) doi:10.2196/14731

KEYWORDS

social media; new media; misinformation; trustworthiness; dissemination; health communication

Introduction

Background

In recent years, there has been significant growth in the uptake of personal communication technologies around the world. This has been largely afforded by the widespread availability of social media (SM) and has been facilitated by the increase in mobile phone ownership. SM has become a valuable tool for communication and it has been utilized in many areas, such as education [1], marketing [2], and health communication [3]. For example, in the field of health communication, the US Centers for Disease Control and Prevention (CDC) [4] and local health departments in the United States [5] have used Twitter to communicate to people during epidemics. Another example is from the United Kingdom and Norway, where health authorities used Twitter to inform their citizens during the West African Ebola outbreak in 2014 and 2015 [6].

The use of SM can improve the nature of health communication as it speeds up the interaction between health care organizations, professionals, and patients [3]. Thus, various SM platforms and apps can play a vital role in health communication and in the promotion of good health [7]. Despite the advantages that SM potentially offers for health communication, it also faces certain challenges. For example, during a health crisis, there is only a limited amount of time for authorities to respond in an efficient way and inform people, while simultaneously helping to eliminate uncertainty on a topic. If this does not occur promptly, it is much more likely that rumors will spread, possibly through SM; when this happens, the negative effects of SM, such as confusion and misinformation, are the probable results [8].

Illustrative examples include the negative consequences experienced by Saudi Arabia and African countries during the Ebola and Middle East respiratory syndrome (MERS) outbreaks. In Saudi Arabia, SM rumors prevented some people from going to emergency departments when they were in an acute condition, resulting in the cancellation of their surgical procedures [9]. In Africa, rumors were found on SM that drinking a huge amount of salt water was a cure for Ebola; it has been reported that this may have caused the deaths of several people [10]. Notably, these misinformation issues seem to affect developing countries more deeply; studies have suggested that 4.5% of Twitter posts (ie, tweets) in the United States are misinformed [11], however, Oyeyemi et al [10] found evidence showing that 50% of tweets in West Africa are misinformed.

Three different studies conducted on health and different types of users reported that Twitter was the preferred platform among health professionals [12], medical students [13], and diabetic patients [14]. In addition, most health-related studies on SM focus on the English-speaking population and the United States [15,16] and not on other cultures. Specifically, Hagg et al [17] reported the absence of literature analyzing SM data for health-related purposes in the Middle East. This is particularly surprising because recent statistics from Statista indicate that Saudi Arabia has the fourth-highest number of Twitter users in the world [18]. Furthermore, when assessing the ratio between Twitter users and populations for each country [19], Saudi Arabia has the highest number of users on Twitter relative to its population. These findings are also supported by other researchers who reported on the elevated prevalence of Twitter usage in Saudi Arabia [20,21]. That being said, only one study—Alnemer et al—has analyzed Saudi health tweets [22].

Given the likely cultural differences between Twitter use in the West and in the Middle East, it seemed important to assess health-related tweets in a Middle Eastern country. Given the prevalence of Twitter use in Saudi Arabia, that seemed like an appropriate country to choose. Hence, this study focuses exclusively on Saudi Arabia.

While a number of tweet characteristics have been assessed for accuracy of the information on SM, the foremost characteristic of interest in this regard has been the source of the tweet. Intuitively, one would anticipate that tweets from health professionals would be more trustworthy; however, this is an open question. A study by Alnemer et al [22] found that 50% of Saudi health professionals' tweets were not evidence based. In addition, this study only includes tweets that were posted by accounts with more than 45,000 followers. The relatively high number of followers suggests that these account holders might be considered opinion leaders in their domain of expertise, which, in this case, is health.

These findings question the accuracy of health professionals' tweets, which is a worrying result considering that people are traditionally more likely to trust users who are physicians, health organizations, and pharmacists [23-27]. While these sources are trusted, evidence shows that they are not necessarily trustworthy (ie, accurate); Alnemer et al [22] suggest that, even if sources are traditionally trusted, there is a high possibility that they include inaccurate (ie, untrustworthy) information.

A few methods and tools for detecting misinformation on SM, particularly during a health crisis, have been proposed and are usually focused on specific topics and diseases (eg, Ebola and Zika) [28,29]. They typically strive to identify the characteristics of misinformation, while neglecting the factors that indicate trustworthy tweets.

In this study, tweets are considered to be trustworthy if they are accurate and are considered to be untrustworthy if they are inaccurate. This position is similar to Yin et al [30], who state that a website is trustworthy if it provides correct information, and information is likely to be true if it is provided by a trustworthy website. Likewise, Zhao et al [31], who developed a topic model to estimate the trustworthiness of the news on Twitter, defined a trustworthy tweet as one that refers to things that really happened. Similarly, this paper considers a tweet as trustworthy if it contains accurate health information and if the process of evaluating the accuracy of tweets is introduced in the methods section.

The perspective we take in our work is to focus on determining the factors that correlate with the trustworthiness of health information tweets as well as the factors that affect the dissemination of those tweets. This work has been undertaken in order to determine how SM might be effectively oriented toward the dissemination of trustworthy health information.

We assess the trustworthiness of tweets originating from traditionally trusted health sources and examine the relationships between attributes (>100) of a tweet and its trustworthiness.

Tweets do not need to be accurate or evidence based to receive traction. In their study, Nastasi et al [32] noted that scientifically inaccurate health tweets were retweeted in the same manner as accurate tweets. That work also indicated the need to study the dissemination metrics of tweets in order to find factors that correlate with high dissemination. Consequently, as a second objective, we will assess the factors that correlate with larger dissemination of health tweets.

Prior Work

SM data from Facebook, Instagram, YouTube, and Twitter have been used to understand people's attitudes and behaviors in sharing and consuming information related to specific health issues, such as vaccinations, abortions, posttraumatic stress, and cancer [33-38]. Facebook has been used by both private health stakeholders and government agencies to engage with the public [39,40]; studies have provided understanding by analyzing Facebook's timelines [40] and health agencies' accounts on Facebook [39].

Although Facebook is the most popular overall platform, the most popular SM platform to study health is Twitter [41-43], as evidenced by studies included in different systematic reviews of topics related to health and social media [44-47]. This focus may be due to the complexity of Facebook data and its unavailability due to privacy restrictions [41].

Most health studies on Twitter collect their data by using specific keywords [28,29,48-52] or from tweets authored by specific health stakeholders, such as health organizations [5,22,53-58]. However, it appears that most studies that analyze Twitter for health using specific keywords have not analyzed the types of users who post the tweets [28,29,48-52]; it is known from other studies that there are different types of users and they share different types of information or hold different attitudes toward specific health issues [59-62]. For example, a number of studies performed tweet extraction using keywords to identify public concerns during the Zika outbreak [48-52]. Besides being limited to a particular outbreak, these studies did not analyze the interplay between public concerns and the types of users and did not address the factors that make a tweet trustworthy.

A notable study of health-related tweets in Arabic was the one performed by Alnemer et al [22], which manually analyzed the accuracy of tweets authored by preselected health accounts. Their results suggested that governmental institutions are more likely to tweet accurate information than are health professionals or other institutions (ie, physicians, dieticians, and nongovernmental and unofficial health institutions). They reported that 80% of the observed governmental institutes' tweets consisted of accurate health information, followed by physicians (60%). However, the overall accuracy of the tweets, over all observed accounts, was 50%. These findings suggest that even if SM users have health expertise, it cannot be taken for granted that their tweets provide accurate health information. This line of investigation can easily be extended to nonhealth users. Alnemer et al [22] did not examine the characteristics of a tweet that may correlate with its trustworthiness.

In terms of trustworthiness, a number of classifiers for health-related tweets in English have been proposed. For example, Ghenai and Mejova [58] proposed a classifier to detect health rumors on Twitter limited to the Zika virus. A limitation of this study is the fact that their classifier was trained on a limited number of rumors, identified as such by information on external non-SM websites. In addition, annotators who labeled the tweets as misinformed were not health experts. In another study, Ghenai and Mejova [29] focused on the detection of users tweeting or propagating misinformation about cancer, excluding social bots and organizational accounts.

However, social bots have also been considered as a possible source for health misinformation on SM [63,64]. For example, Allem et al [64] analyzed tweets in regard to e-cigarette discussions and found that social bots may support misinformation on SM in regard to e-cigarette cessation. In their study, they emphasized the importance of distinguishing between social bots and real users. Similar to Allem et al [64], Broniatowski [63] analyzed tweets specifically to understand how bots promote online health content in regard to vaccine-related messages and found that social bots were one of the possible sources for antivaccine advice on SM.

As suggested above, in terms of social bots and misinformation on SM, previous research has emphasized the importance of distinguishing social bots from real users, particularly when the intent is to assess views held by users who are not bots [64,65]. However, this is not an easy task, as some social bots might mimic user behavior [66]. Social bots might introduce themselves as individual accounts with locations and photos for their profiles [64,66]. Furthermore, some organizations use social bots to disseminate information, which makes it hard to distinguish between it being the *opinions of the bot* or that of their organization and classifying according to their organization might be difficult [67,68]. In the work presented here, users were classified as they introduced themselves and it was that classification that was analyzed in terms of the accuracy of information they portrayed.

Kalyanam et al [28] analyzed the association between hashtags and the credibility of tweets related to the Ebola outbreak. They defined credible tweets as those with hashtags that indicated origin from well-known governmental agencies or other authoritative sources (eg, #cdc or #cnn) and speculative tweets as those with hashtags that indicated the spread of fear, rumor, scam, or humor. It was determined that almost 25% of the analyzed tweets were speculative. Their findings suggested that verified users were more likely to interact with credible hashtags; on average, the number of followers for accounts that posted tweets with credible hashtags was 7000, compared to 2700 for accounts that posted tweets with speculative hashtags in their dataset. Kalyanam et al relied on hashtags, without evaluation of the information carried by the tweets. That is, it is unclear whether tweets classified as credible really contain accurate or trustworthy information.

In terms of identifying influential users on SM, Albalawi and Sixsmith [69] applied six different tools to identify the most

influential Twitter users in Saudi Arabia. First, they used the apps Tweepar and SocialBaker, which reveal the most influential users per country on Twitter. With the influential users being identified, they collated four Twitter influence scores via the following: Social Authority by Moz [70], PeerIndex [71], Kred [72], and Klout [73]. However, within the scope of their study, they did not consider health in isolation and they did not analyze the accuracy of the tweets.

Wong et al [5] analyzed tweets sent by 287 local health departments (LHDs) during the Ebola epidemic in the United States. They found that 70% of the LHDs tweeted at least once about Ebola and that Twitter had become a frequent tool used by LHDs during this particular epidemic. Regarding the dissemination of tweets, one of their findings was that the presence of hashtags and links was highly correlated with the messages being retweeted. Similarly, Suh et al [74] also reported that tweets containing hashtags and links were more likely to be retweeted. They did not consider the impact of the type of users on retweeting. Furthermore, the analyzed tweets were on randomly selected topics, unrelated to the health care domain.

The research results summarized in this section suggest that there is a lack of comprehensive studies on the accuracy of health-related tweets on SM. Specifically, to the best of our knowledge, there are no results that determine the factors that make a health-related tweet trustworthy in general, besides it being authored by a credible institution. In addition, there are no studies on the general factors that affect the dissemination of trustworthy health-related tweets, besides the credibility of the author and the presence of hashtags and links. We believe that the identification of such factors may help health organizations in better disseminating trustworthy information during an outbreak-related health crisis. In particular, there is a lack of studies on health-related tweets in Arabic; this makes such a study a priority, considering the high popularity of Twitter in the Arab world.

Hence, the work presented in this paper addresses the following questions:

1. How can the trustworthiness of health care stakeholders' tweets be identified from the tweets' features?
    a. What proportion of trustworthy health-related tweets come from the following sources: health professionals, health organizations, and authorities?
    b. What are the other characteristics associated with trustworthy health-related tweets?
2. What are the factors that contribute to the wider dissemination of health care-related tweets that could possibly be used to make accurate health information dominant over other related information on SM?
    a. Does the trustworthy nature of health-related tweets increase their dissemination?
    b. What other factors contribute to the dissemination of health care-related tweets?

The structure of this paper is as follows. First, the paper introduces the empirical design for identifying the factors affecting the trustworthiness and dissemination of health-related tweets. Second, it presents the results of our work and highlights, in particular, the findings that correspond to the research questions listed above. Finally, this paper ends by discussing the future directions of our research and the possible implications of the results.

Methods

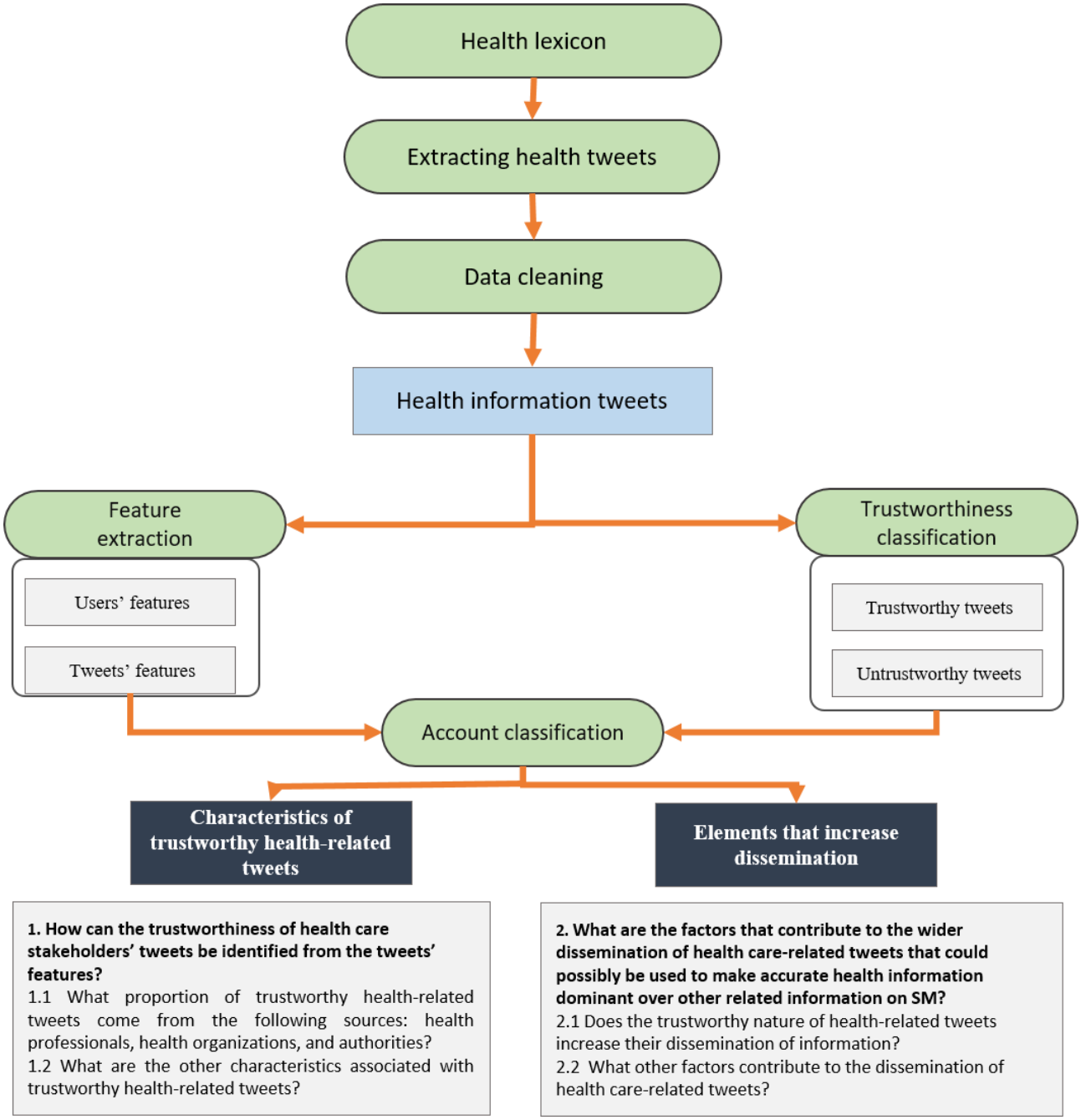
Overview

This work utilized a standard text analytics methodology, which incorporated the following steps:

1. We developed a health lexicon using two different methods.
2. By means of this health lexicon, we extracted health tweets using the Twitter application programming interface (API).
3. From the remaining tweets, we manually refined tweets related to health using two annotators.
4. Medical professionals manually labeled the remaining tweets as either accurate or inaccurate.
5. We extracted features from the labeled dataset. These included attributes of the tweets as well as attributes of the user profiles for the users who authored the tweets. In this paper, only aggregated user data is presented for ethical reasons.
6. We analyzed the labeled dataset to provide preliminary answers to the research questions outlined in the previous section.

The outlined methodology is presented in Figure 1. The sections that follow explain each step in detail.

Figure 1. Overview of the research methodology. SM: social media.



Construction of the Health Lexicon

In order to identify health-related tweets, a health lexicon was created. It is important to note that the incorrect selection of indicative health keywords could bias the results [75]. Therefore, two separate methods of analysis were utilized to generate this lexicon. The first method consisted of asking three medical doctors with active Twitter accounts to provide 100 health-related words. The doctors we asked are skilled in different disciplines and differ from each other in their age, background, and gender. They were asked to provide an initial list of health-related words that they think would be used in tweets related to health. The second method involved the usage of 110 health care keywords that were independently identified by an annotator with a college degree in linguistics. As this study concentrates on Saudi Arabia, the annotator identified these keywords by examining a set of tweets with geolocation that indicated a Saudi Arabian origin, although people who enable geolocation are likely to represent a specific demographic group [76,77]. As such, the words chosen by the annotator might have limited generalizability with respect to the wider demographic group. Thus, the annotator also reviewed health-related accounts and hashtags to identify different health-related words. A complete list of keywords is attached in Multimedia Appendix 1.

These two methods were combined in an attempt to construct a health lexicon that is as unbiased as possible.

Data Cleaning

Using the lexicon developed as described in the previous section, it was then possible to extract Twitter data for the main part of

our work. The Twitter API does not allow users to extract tweets more than a week old [78]. To reduce the impact of this limitation, it was decided that we would extract two datasets. The first dataset was extracted on May 18, 2018, and the second dataset on August 7, 2018. Table 1 describes the characteristics of each dataset.

Table 1. Characteristics of the datasets.

Characteristics	First dataset, n (%)	Second dataset, n (%)
Total tweets	209,345 (100)	196,670 (100)
Original tweets	57,794 (27.61)	39,454 (20.06)
Reply tweets	28,329 (13.53)	32,470 (16.51)
Retweeted tweets	123,222 (58.86)	124,746 (63.43)

Reply tweets were difficult to evaluate by annotators, due to a lack of contextual information. As a result, these tweets were removed. It was also necessary to remove all retweeted tweets due to redundancy.

The tweet extraction process resulted in an accumulation of 97,248 tweets: 57,794 from the first dataset and 39,454 from the second. Using the random method in Python, version 3.6.5 (Python Software Foundation), a sample of 2800 tweets was selected from the set for use in a prototype study. Even though the lexicon suggested that all of the 97,248 tweets were health related, a manual examination of the tweets showed that this was not necessarily so. Thus, two annotators were employed to filter out tweets unrelated to health from the sample of 2800 tweets. The guidelines for the annotators were as follows:

1. Tweets that describe any function of the body, such as enzymes, organs, or diseases, should be retained.
2. Tweets that give advice or information about supplements, drugs, physical activity, or food and link it to people's health, such as how vitamins, food, and drugs affect people's health, should be retained.

Based on an internal discussion among the three authors of this study, these guidelines were derived from Bobicey and Sokolova's ontology for personal health information [79]. The terms were derived from concepts in their ontology and were considered by all three authors as the most indicative of health-related material.

Each annotator labeled 60% of the tweets, with 10% of the tweets (n=280) labeled by both annotators in order to check the reliability of the analysis. The Cohen kappa statistic for interrater reliability [80] was then calculated, resulting in a value of .872, which indicates excellent agreement between the annotators.

Out of the 2800 tweets, only 552 tweets (19.71%) were labeled as health related. Out of these 552 tweets, 180 tweets (32.6%) originating from the first dataset were selected, in addition to 120 tweets (21.7%) originating from the second dataset. By doing so, it was possible for us to retain the proportion of tweets in the originally collected datasets.

Trustworthiness Classification

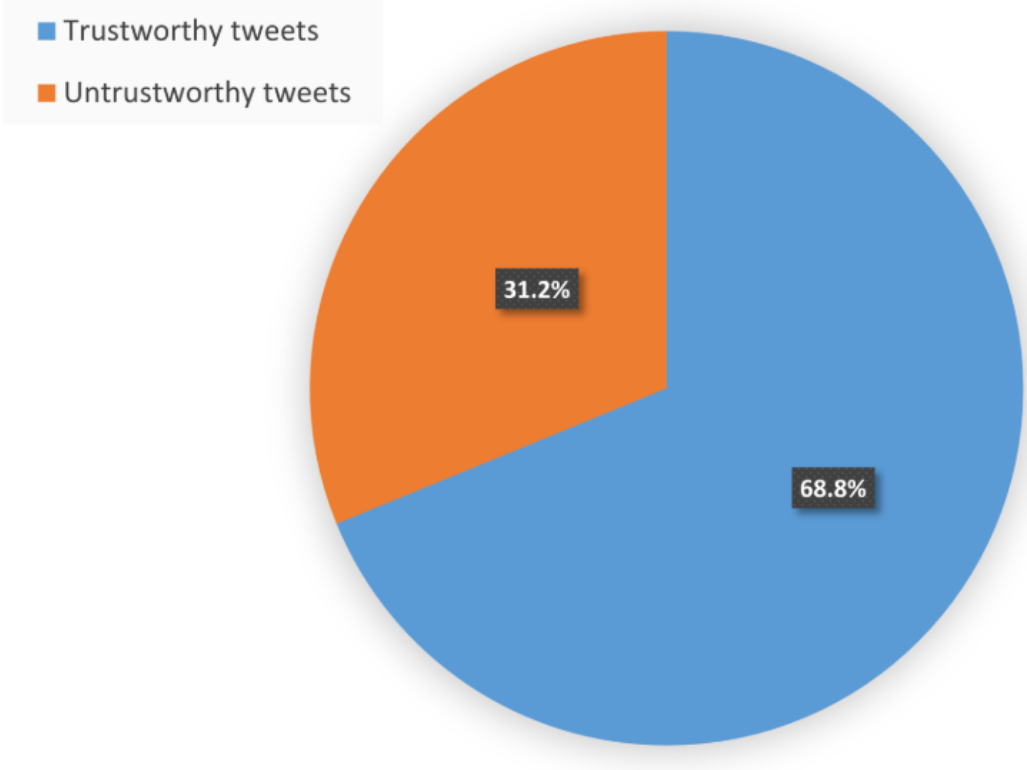
Once the previous processes had been completed, 10 medical doctors were asked to manually classify the tweets into the following categories:

1. Accurate health information.
2. Inaccurate health information.
3. Not sure about the accuracy.

The *not sure* option was given to the doctors to avoid forcing them to make a decision on tweets if they did not have enough relevant health knowledge to accurately evaluate them or if the tweets were ambiguous.

A total of 10 Google forms were created, each containing 30 tweets. A link to each form was sent to two doctors by email. In order to achieve high reliability with respect to the accuracy and inaccuracy, we excluded any tweets that a doctor labeled as *not sure*. In addition, we excluded any tweets where the two doctors coding them disagreed on their accuracy. This resulted in 109 labeled tweets in the spreadsheet, 75 (68.8%) of which were labeled as trustworthy (ie, tweets that both doctors labeled as accurate) and 34 of (31.2%) of which were labeled as untrustworthy (ie, tweets that both doctors labeled as inaccurate). The information was then transferred to a spreadsheet and analyzed using Python for descriptive statistics. For the statistical tests, we used the R package, version 3.4.0 (The R Foundation) [81]. The output of this process is illustrated in Figure 2.

Figure 2. Proportion of trustworthy (n=75) and untrustworthy (n=34) tweets in the sample of 109 labeled, health-related tweets.



Feature Extraction

Overview

To answer the research questions, it was necessary to determine the features of the tweets as well as their level of trustworthiness. The features of the tweets were categorized into two types: tweet features and user features.

These features, whether on the user level or the tweet level, may provide data that is useful in classifying the types of users or identifying the credibility of the tweets for different topics [82-86].

Tweet Features

Tweet features were extracted directly from the tweets. These included whether or not the tweet was retweeted (ie, as a dissemination measure) as well as various linguistic characteristics of the tweets, such as the number of words and the number of characters in each tweet. Tweet features also included other properties of the tweets, such as URLs contained in the tweet, the time of the tweets, and hashtags. Most of these features have been used by other researchers; however, the linguistic features identified in most other studies were analyzed for Latin-derived words. Thus, more features were added after reviewing literature related to Arabic natural language processing. These features are as follows:

1. Tashkeel: the presence of a tashkeel in the tweet. The tashkeel is a special Arabic character written in the text to represent missing vowels [87].
2. Kashida: the presence of a kashida in the tweet. The kashida is a decorative element in Arabic writing used for justifying the text [88].

In addition, similar to Castillo et al [82], we examined different types of punctuation marks in tweets and their relationship to information credibility. This was based on our insight that people who used punctuation in their tweets appeared more thorough and that this might be associated with greater accuracy and dissemination.

A complete list of tweet features is provided in Multimedia Appendix 2.

User Features

Overview

Tweets come with metadata that provide basic profile information about the user who posted the tweet, such as screen name, number of friends, number of followers, favorite count, retweet count, and age of the account. In addition, there are cumulative tweeting characteristics for each user. To derive this data, we used the Twitter API to extract another 200 tweets per author by the authors of the 109 labeled tweets. The tweet number of 200 has been suggested as sufficient for extracting user features [83,89]. The user features were categorized into four groups, which will be described in the following sections. A complete list of these features is provided in Multimedia Appendix 2.

Activity and Connectedness Features

These features include metrics that measure how active the user is, such as how often the user replies to other users or how often the user retweets [86].

User Linguistic Features

These features include measures of how often the user uses unique hashtags, the average number of hashtags used in tweets, and the mean number of words in the user's tweets. Such linguistic features have been used in other studies to assess information credibility on Twitter [82] and to classify users on SM [83,90].

User Time Features

These features deal with the temporal aspect of the user's tweets. For example, previous research examined the day of the week when tweets were posted to determine if the day is linked to the credibility of news [82]. In the study presented here, more features were added, such as the preferred time of the day—morning, evening, and night—for users to tweet and whether the tweet was posted during weekdays or weekends [91].

User Popularity Features

These features indicate the popularity of the users, such as the number of followers per user as well as how often users' tweets were retweeted [86].

Account Classification

Accounts were classified according to the following criteria [53,59,61,62,69,92]:

1. Does the account holder have a health background?
2. Is the account that of an individual or nonindividual?
3. If the account holder is a nonindividual, does the account represent a physical authority or is it an exclusively SM-based account?

Hence, by the end of this process, it was expected that the following categories of user accounts could be analyzed:

1. Individual health accounts.
2. Individual nonhealth accounts.
3. Health organization accounts.
4. Nonhealth organization accounts.
5. Exclusively SM-based accounts.
6. Users whose profiles cannot be extracted and, therefore, remain unknown.

Two annotators classified the types of tweets in the dataset based on author accounts and disagreed on seven users. They met to explain their opinions to each other. Finally, after a discussion they agreed on the categories of five of these seven users. An expert in health communication on SM was consulted—a surgeon with a PhD in Health Promotion in New Media—to classify the final two users. In addition, there were about 10 accounts for which profile data could not be extracted; these accounts were classified as *unknown*.

Data Analysis

Lancaster et al [93] recommended that the execution of a pilot study should primarily rely on descriptive and distribution statistics as results. For the continuous variables, it was decided to present the median number, as the median is not affected by the outliers. For the categorical variables, it was decided to employ a statistical test to determine preliminary results. As our data is nonparametric, the Mann-Whitney-Wilcoxon (MWW) test was utilized to establish statistical significance. The MWW test is considered appropriate for nonparametric data and for when the two samples are from different populations (ie, in our case, accurate tweets and inaccurate tweets) and of different sizes [94].

For categorical variables, the Fisher exact test was also used, which is suitable for a small sample of less than 1000 [94]. In the study presented here, there were two categorical variables with more than two values: the type of author and the times at which tweets were posted. Typically, the Fisher exact test does not test for statistical significance for a contingency table larger than 2×2; however, in R it is possible to calculate the *P* value for larger contingency tables, hence R was used here. The mechanism that allows for the calculation of the *P* value is based on the work of Mehat and Petal [95] and Clarkson et al [96].

Results

In this section, the most promising results are presented, which indicate factors that may demonstrate the trustworthiness and untrustworthiness of health-related tweets. Overall, more than 100 tweet-level and user-level features in a dataset of 109 tweets were explored; these tweets were labeled as either accurate and trustworthy or inaccurate and untrustworthy.

An initial analysis of the tweet-level features indicates that trustworthy health tweets were significantly more likely to have an author that is a member of a list (ie, a curated group of Twitter accounts) (P=.05). Although not significant, trustworthy tweets seemed more likely to be favorited by others (P=.06).

In contrast, Table 2 suggests that untrustworthy health tweets were more likely to have URLs embedded in them (P=.03). Specifically, 24% (8/34) of the untrustworthy tweets had URLs, compared to 8% (6/75) of trustworthy tweets. A total of 4 out of 8 (50%) of the URLs cited in inaccurate tweets referred to news websites, while 2 out of 8 (25%) referred to blogs.

Table 2. Most promising tweet features to help distinguish the accuracy level of tweets.

Metric	Description	Trustworthy tweets (N=75), n (%)	Untrustworthy tweets (N=34), n (%)	*P* value[a]
URLs	The tweet contains a URL	6 (8)	8 (24)	.03
Listed	The author is listed	53 (71)	17 (50)	.051
Favorited	The tweet is favorited	46 (61)	14 (41)	.06
Hashtags	The tweet contains a hashtag	20 (27)	4 (12)	.13
Tashkeel	The tweet contains a tashkeel	22 (29)	5 (15)	.15
Exclamation mark	The tweet contains "!"	3 (4)	4 (12)	.20
Semicolon	The tweet contains ";"	42 (56)	14 (41)	.21
Retweeted	The tweet was retweeted	43 (57)	15 (44)	.22
Kashida	The tweet contains a kashida	6 (31)	8 (24)	.49

[a]*P* values were calculated using the Fisher exact test with $P \leq .05$ indicating statistical significance.

The user-level features were also analyzed (see Table 3). The analysis revealed the worrying trend that users who had a lower *number of followees* (F3) were more likely to tweet accurate health tweets ($P<.001$). More encouragingly, the popularity measure of *number of times that author's tweets are favorited* (FT2) was associated with trustworthiness, suggesting that accurate tweets were recognized as such. Interestingly, authors who tended to retweet tweets that had hashtags (RMH5) also tended to tweet trustworthy tweets, although the *P* value is not quite significant in that case ($P=.06$).

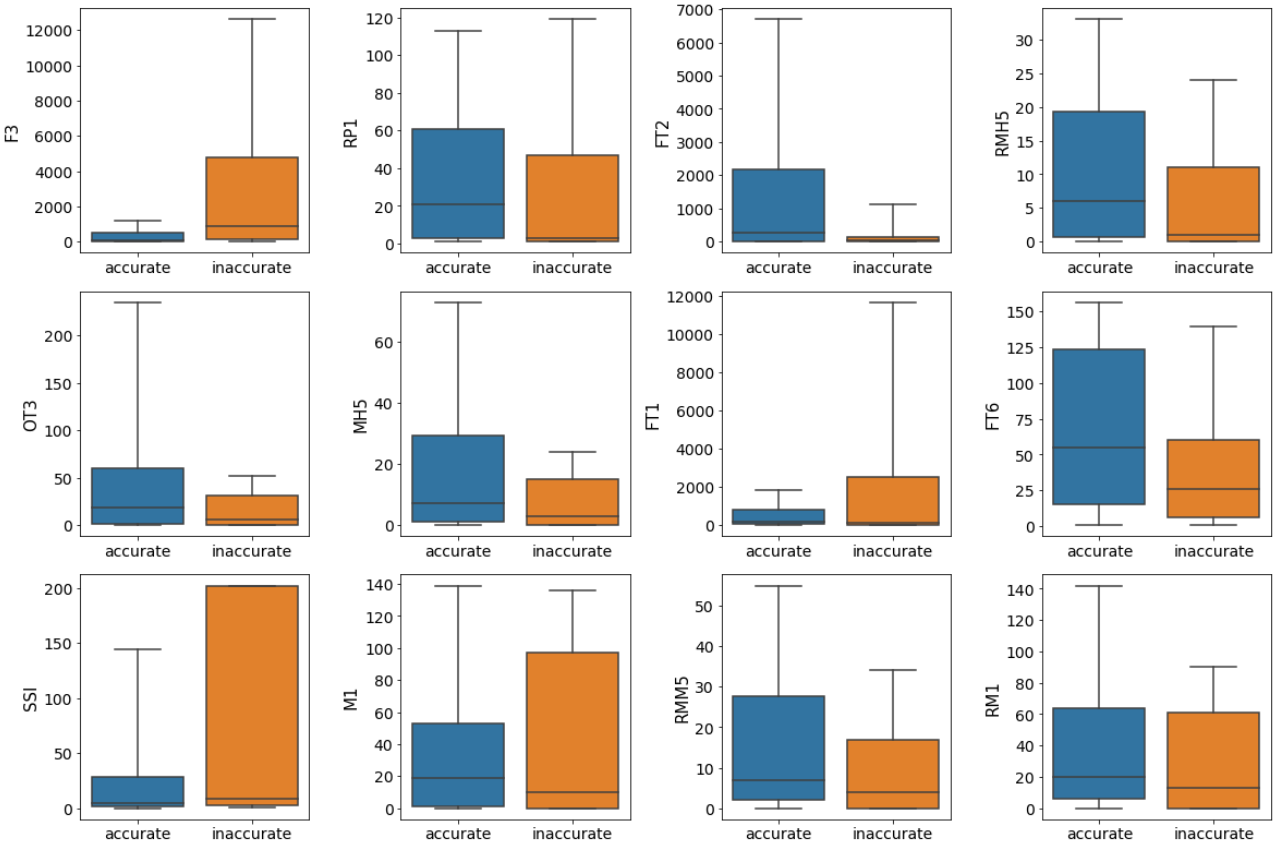
Table 3. Metrics for users who tweet accurate information versus users who tweet inaccurate information.

Metric	Description	Trustworthy tweets, median (SD)	Untrustworthy tweets, median (SD)	*P* value[a]
F3	Followees count	75.5 (2517)	891 (12,952)	<.001
FT2	Number of times author's tweets are favorited (ie, favorite-author tags)	281.5 (76,054)	46 (2454)	.01
RMH5	Unique hashtag count in tweets that were retweeted by the author	6 (24.2)	1 (26.6)	.06
OT3	Number of hashtags in the author's tweets	18 (144.4)	6 (64.5)	.09
RM1	Number of retweeted tweets by the author where the user mentioned other users	20 (76.3)	13 (49.7)	.09
FT6	Number of original tweets posted by the author that are favorited	55 (60.5)	26 (58.7)	.11
SSI	Ratio of original tweets posted by the author to tweets retweeted by the author	4.51 (61.6)	8.94 (93.8)	.11
MH5	Unique keyword count in hashtags set in original tweets posted by the author	7 (46.1)	3 (22.9)	.14
RP1	Number of reply-to tweets posted by the author	21 (47)	3 (54.1)	.20
RMM5	Unique mentions in retweeted tweets by the author	7 (26.1)	4 (24)	.43
M1	Number of tweets where the author mentioned other users	18.5 (61.6)	10 (66.6)	.51
FT1	Number of tweets favorited by the author	179 (1202.6)	83 (14,073)	.84

[a]*P* values were calculated using the Mann-Whitney-Wilcoxon (MWW) test with $P \leq .05$ indicating statistical significance.

Figure 3 shows the boxplots for the user features, with outliers beyond 90% of the data excluded. This figure suggests that there may be other metrics associated with trustworthiness given a larger dataset, such as the *number of tweets favorited by the author* (FT1), the *number of original tweets posted by the author that are favorited* (FT6), and the *ratio of original tweets posted by the author to tweets retweeted by the author* (SSI).

Figure 3. Boxplots for the features most closely correlating with the trustworthiness of tweets; outliers outside the 90th percentile were excluded. F3: number of followees; FT1: number of tweets favorited by the author; FT2: number of times that author's tweets are favorited; FT6: number of original tweets posted by the author that are favorited; M1: number of tweets where the author mentioned other users; MH5: unique keyword count in hashtags set in original tweets posted by the author; OT3: number of hashtags in the author's tweets; RM1: number of retweeted tweets by the author where the user mentioned other users; RMH5: hashtag count in tweets that were retweeted by the author; RMM5: unique mentions in retweeted tweets by the author; RP1: number of reply-to tweets posted by the author; SSI: ratio of original tweets posted by the author to tweets retweeted by the author.



As per Figure 4, regarding the timing metrics, it appears that the majority of the trustworthy tweets were posted in the morning, while most untrustworthy tweets were posted either in the evening or at night. The associated Fisher test is borderline, with a *P* value of .06. However, when comparing only two groups—the morning tweets and the night tweets—statistical significance is achieved with a *P* value of .04. This result aligns with our insight that professional tweets are posted during the daytime and more informal tweets are posted at nighttime.

As can be seen in Figure 5, sources classified as organization accounts were more likely to tweet accurate information, followed by sources traditionally considered as trusted users (ie, health professionals, health authorities, and health organization accounts). However, there were only 10 tweets from organization accounts, one of which was from an organization unrelated to health care. Overall, traditionally trusted users are considered the most trustworthy source in the dataset, as there were 34 tweets from them and only 4 (12%) were considered inaccurate. The least trustworthy category in the dataset includes the users with exclusively SM-based accounts.

When the backgrounds of individual users were considered, those with a background in health care (ie, health professionals) seemed to tweet accurate health information and were less likely to tweet inaccurate health information (see Figures 6 and 7); they posted 29 out of 109 tweets (26.6%). Individuals from a non-health care background were core players in terms of the volume of tweets, with 18 out of 30 (60%) of their tweets labeled as trustworthy. Overall, however, they authored 12 out of 34 (35%) of the untrustworthy tweets and only 18 out of 75 (24%) of the trustworthy tweets.

Another interesting finding is that exclusively SM-based accounts seemed to tweet much more inaccurate health information than other account types. Therefore, to summarize, trusted health accounts (ie, health care organizations and professional health care individuals) were more likely to tweet trustworthy health information than were other types of accounts. Individual users were the main players in terms of the volume of health information on SM, but the high volume did not correlate with trustworthiness. The Fisher exact test indicated statistical significance between the type of the user who posted a tweet and the accuracy of the tweet (*P*=.04).

Figure 4. Distribution of accurate and inaccurate tweets per time of day in three categories: morning (6 am-2 pm), evening (2 pm-10 pm), and night (10 pm-6 am).

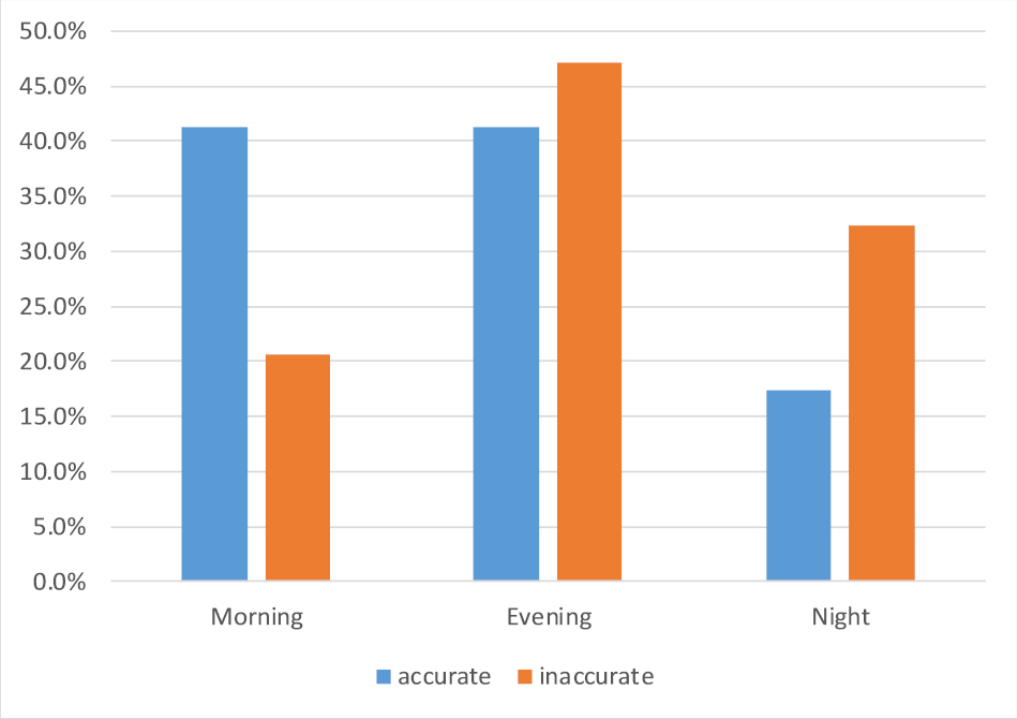


Figure 5. Accuracy of the tweets posted by each author type. SM: social media.

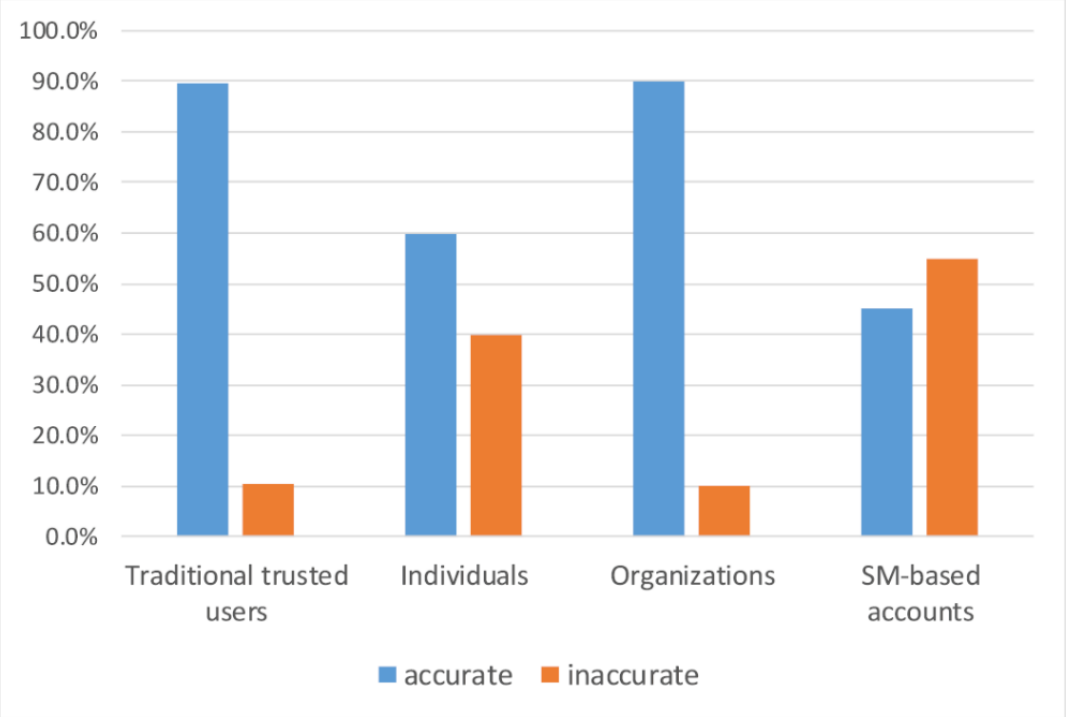


Figure 6. Distribution of the authors of accurate (ie, trustworthy) health-related tweets. esm: exclusively social media-based accounts; ho: health organization accounts; ih: individual health accounts; inh: individual nonhealth accounts; nho: nonhealth organization accounts; unk: users whose profile cannot be extracted and, therefore, remains unknown.

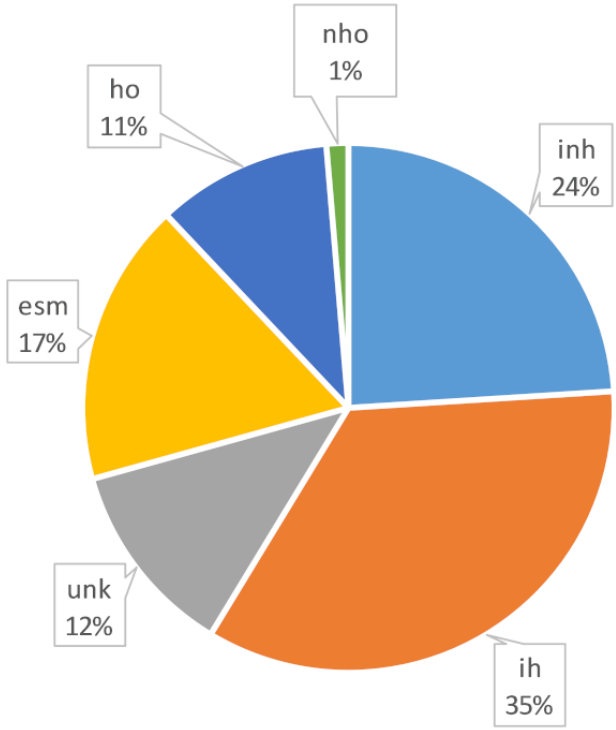
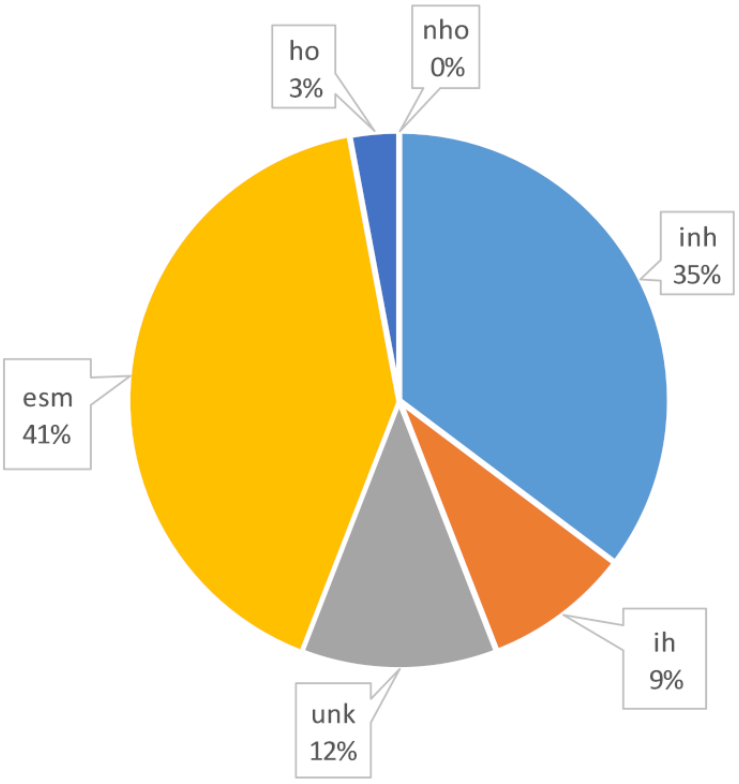


Figure 7. Distribution of the authors of inaccurate (ie, untrustworthy) health-related tweets. esm: exclusively social media-based accounts; ho: health organization accounts; ih: individual health accounts; inh: individual nonhealth accounts; nho: nonhealth organization accounts; unk: users whose profile cannot be extracted and, therefore, remains unknown.



To address the question of whether the trustworthiness of tweets increases their dissemination, a retweeting metric was examined. According to Suh et al [74], retweeting is the key mechanism for information dissemination on Twitter. The results of our study indicate that trustworthy health information was slightly more likely to be retweeted than inaccurate health information,

but the difference was not significant. Specifically, 43 out of 75 (57%) of the accurate tweets were retweeted compared to 15 out of 34 (44%) of the inaccurate tweets being retweeted. This is in line with the findings in other studies, which found that health information does not need to be accurate in order to be disseminated [32]. However, when accurate health-related tweets were retweeted, they were more likely to be further retweeted than inaccurate health-related tweets, as shown in Figure 8.

The provisional findings suggest that tweets with embedded commas, listed authors, or marked as favorite were also associated with dissemination (P<.001), as indicated in Table 4, which also lists a few other factors we examined. The only factor that was clearly counter-indicative of dissemination was the presence of a URL in the tweet, with only 5.2% of these tweets being retweeted.

Table 5 shows statistical significance for most of the popularity metrics.

Figure 8. Retweeted counts of accurate and trustworthy versus inaccurate tweets in our dataset; outliers outside the 90th percentile were excluded (P=.043).

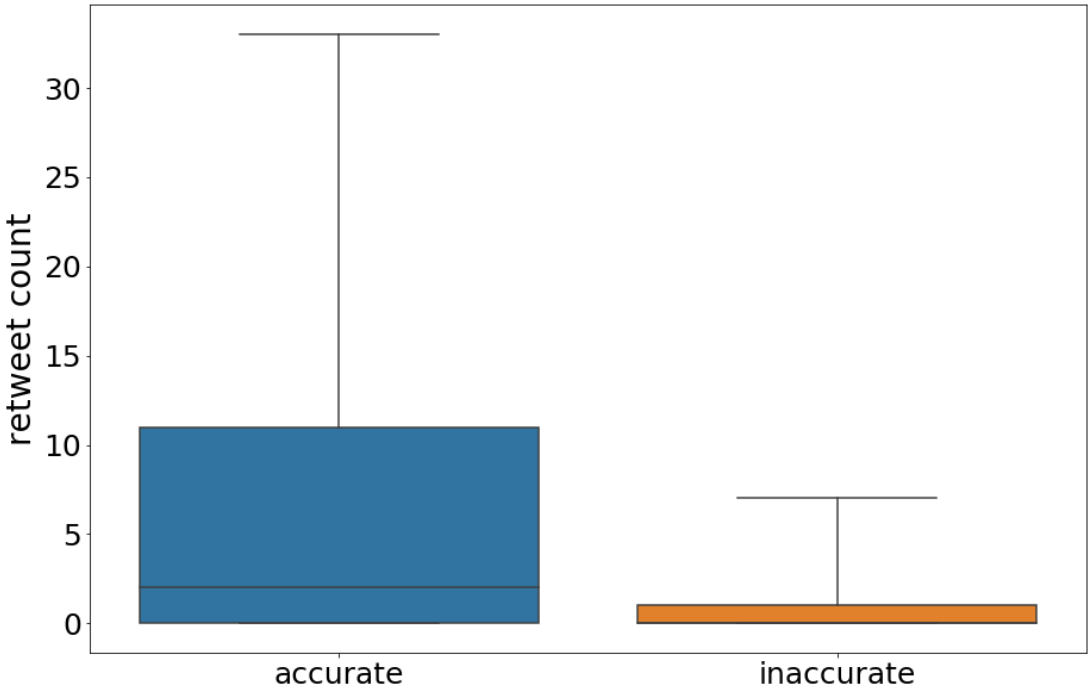


Table 4. Features for retweeted tweets versus unretweeted tweets.

Metric	Description	Retweeted tweets (N=58), n (%)	Unretweeted tweets (N=51), n (%)	*P* value[a]
Listed	The author is listed	50 (86)	20 (39)	<.001
Favorited	The tweet is favorited	51 (88)	9 (18)	<.001
Comma	The tweet contains a comma	25 (43)	9 (18)	.01
Kashida	The tweet contains a kashida	22 (38)	9 (18)	.02
URLs	The tweet contains a URL	3 (5)	11 (22)	.02
Tashkeel	The tweet contains a tashkeel	19 (33)	8 (16)	.04
Tweet accuracy	The tweet is accurate	43 (74)	32 (63)	.22
Hashtags	The tweet contains a hashtag	14 (24)	10 (20)	.64

[a]*P* values were calculated using the Fisher exact test with P<.05 indicating statistical significance.

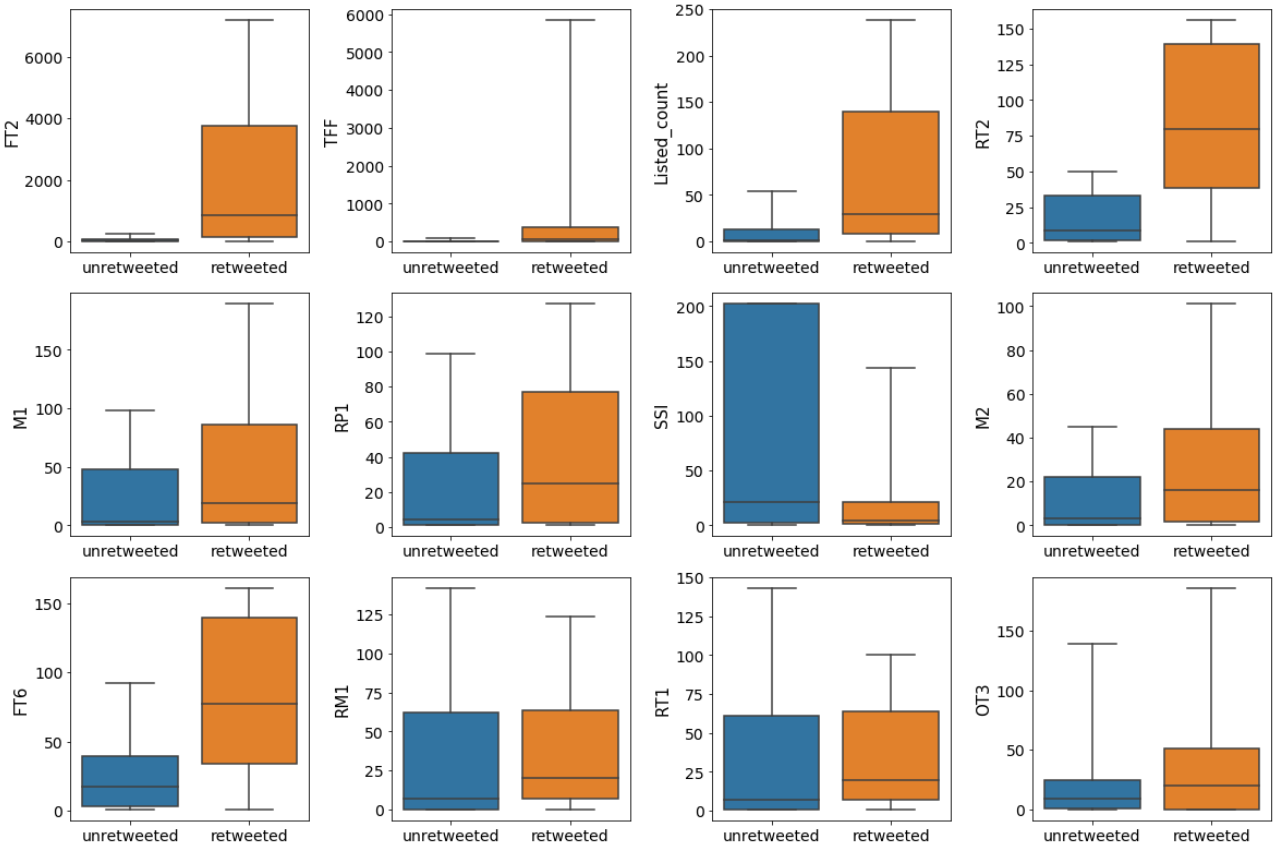
Table 5. Metrics for users who tweeted accurate information versus users whose tweets were retweeted.

Metric	Description	Retweeted tweets, median (SD)	Unretweeted tweets, median (SD)	*P* value[a]
FT2	Number of times that author's tweets are favorited	857.5 (8366)	25 (121)	<.001
TFF	Ratio of followers to followees (F1/F3)	81.6 (18,950)	1.09 (22,728)	<.001
Listed count	Number of lists where the user is member	29.5 (912.1)	1 (99)	<.001
RT2	Number of author's tweets retweeted by other users	79.5 (59.8)	9 (25.5)	<.001
FT6	Number of original tweets posted by the author that are favorited	77 (61.5)	17 (42.3)	<.001
M2	Number of unique users mentioned by the author	16 (43.1)	3 (23.5)	.052
RP1	Number of reply-to tweets posted by the author	24.5 (53)	4 (42.2)	.08
SSI	Ratio of original tweets posted by the author to tweets retweeted by the author	3.83 (63.4)	21.11 (87.6)	.11
M1	Number of tweets where the author mentioned other users	18.5 (67.4)	3 (47.6)	.12
RM1	Number of retweeted tweets by the author where the user mentioned other users	20 (78.3)	7 (56.2)	.12
RT1	Number of tweets that the author retweeted	19.5 (49.1)	7 (54.4)	.14
OT3	Number of hashtags in the author's tweets	20 (142.5)	9 (103)	.26

[a]The *P* value was calculated using the Mann-Whitney-Wilcoxon (MWW) test with $P<.05$ indicating statistical significance.

The user features *ratio of followers to followees* (TFF), FT2, *number of author's tweets retweeted by other users* (RT2), FT6, and listed count appear to have strong associations with dissemination, as shown in Figure 9. This comes as no surprise, as these features are typically considered to be popularity metrics. Furthermore, the results, as indicated in Table 5, show that users who mention other users/posters of tweets (M2) are more likely to have their tweets retweeted.

Figure 9. Boxplots for the features most closely correlated to dissemination; outliers outside the 90th percentile were excluded. FT2: number of times that author's tweets are favorited; FT6: number of original tweets posted by the author that are favorited; M1: number of tweets where the author mentioned other users; M2: number of unique users mentioned by the author; OT3: number of hashtags in the author's tweets; RM1: number of retweeted tweets by the author where the user mentioned other users; RP1: number of reply-to tweets posted by the author; RT1: number of tweets that the author retweeted; RT2: number of author's tweets retweeted by other users; SSI: ratio of original tweets posted by the author to tweets retweeted by the author; TFF: ratio of followers to followees.



Discussion

Principal Findings

In terms of individual health tweets, the results of this study do not agree with those of Alnemer et al [22], in that they suggested that 50% of the tweets were not evidence based. However, Alnemer et al labeled every tweet in their dataset as either accurate or inaccurate, while in this study, tweets for which not all annotators agreed were excluded. This difference in annotation may explain the differences in the results. Nevertheless, this contradiction indicates the importance of conducting future research to explain these distinctions.

In this preliminary analysis, a group of users linked to the accuracy of the health information was identified, indicating that trusted health users are more likely to tweet trustworthy health information than inaccurate health information. This association is supported by the findings of Medlock et al [26].

Nevertheless, a high proportion of tweets from individuals with no health background were also found to be accurate. This observation suggests the existence of a subgroup of trustworthy SM accounts. The isolation of such a subgroup might be possible through the identification of other characteristics.

Both Wong et al [5] and Suh et al [74] reported that interacting with hashtags was linked to dissemination, while this study provided no clear evidence of such a relationship. Instead, we found that the more a user interacted with other users (ie, *number of times that author's tweets are favorited* [FT2]), the more likely it was that their tweets were accurate. This finding suggests that trustworthy users have more influence than other users, as FT2 is considered to be an influence metric [86].

Interestingly, the data revealed that most of the accurate health tweets were posted in the morning, while most of the inaccurate tweets were posted at night. This disparity may occur because health professionals may tweet accurate health information while they are at work, possibly as part of their job, while less-trustworthy tweets are more likely posted at night when nonprofessionals are more likely to give an opinion.

In addition, there is no clear answer as to whether trustworthiness is linked to dissemination, because trustworthy tweets were only slightly more likely to be retweeted. However, when considering the retweet count, accurate tweets were more likely to be retweeted more frequently, as shown in Figure 8. These preliminary results suggest that there is an association between trustworthiness and the ultimate dissemination of the tweets.

Similar findings were also noted by Kalyanam et al [28]; it raises the question as to why trustworthy tweets are more likely to be retweeted more frequently once they are retweeted. One interpretation might be that there are thresholds for followers

that can be exceeded and once they are exceeded, the author might have a certain leverage for their tweets to be retweeted more [97]. However, neither this study nor that of Kalyanam et al looked at self-retweeting specifically. This practice is known to be common in microblogs, such as in circumstances where users retweet to win prizes. Surprisingly, tweets with embedded commas and kashidas were retweeted more, suggesting that correct punctuation may be perceived as a sign of accuracy.

Moreover, some tweet metrics appeared to be linked to both dissemination and trustworthiness; for instance, tweets that embedded the tashkeel were more likely to be retweeted and indicated a trend of possibly being more trustworthy, while tweets that embedded URLs were less likely to be retweeted and trustworthy. These findings contradict those of Wong et al [5], who analyzed specific health accounts and found that URLs in tweets were associated with dissemination.

In regard to the source of the URLs cited in inaccurate tweets, our findings indicated that news websites were the most cited (50%). This is in line with Ghenai and Mejova [58], who found news websites were the most cited sources in inaccurate tweets (39% of the URLs).

Our findings, as shown in Tables 2 and 4, suggested that the language characteristics of the tweets might be associated with both dissemination and trustworthiness. At a high level, this may suggest that the style in which tweets are written is also linked to dissemination and trustworthiness. At a low level, some of these features are language specific; for example, the tashkeel is used in Arabic but does not exist in Latin languages. This specificity indicates the need to take language type into account when designing any future study. The tashkeel was not tested for significance in trustworthy tweets; as *P* was equal to .10, these results should be considered for future studies.

In future work, we will seek to develop a machine learning model for classifying health-related tweets as either trustworthy or not trustworthy. To do this, we aim to employ a larger dataset and to evaluate the usefulness of a larger set of features as predictors. Some of these features may include measures for the linguistic ability of a user. The extraction of additional features may also require the development of additional machine learning models, such as models for topic detection in order to measure, for example, how often a user tweets about health.

Internal Validity

Due to the limited data size, this study provided preliminary findings on the relationship between the variables studied. In addition, this study did not establish any causal relationship between variables, only correlations.

External Validity

The selection of data in this study was limited to health-related tweets in the Arabic language on Twitter. In addition, we collected tweets from a limited time period: two periods of 7 days. Ultimately, we analyzed a small number of tweets and, as a result, we cannot presume generalization for the findings of this study.

We did not include any reply tweets in the analysis; therefore, our results cannot be generalized to interaction-type communication on SM (eg, if the user posts a direct health question to another user). Our results only refer to initial tweets.

Although we used two methods in developing a lexicon, we cannot claim that the lexicon is totally representative of the population. Secondly, in this study we only studied 109 of the 300 tweets labeled by doctors. We excluded tweets where one of the doctors was unsure of the accuracy of the tweet or where there was disagreement between the two doctors regarding the accuracy of the tweet. These measures certainly excluded some health tweets and, more worryingly, may have thus excluded a class of health tweets that were not studied. However, the protocol did heighten the quality of the data in terms of its accuracy. In addition, all 20 doctors who participated in this study as annotators were from the same country, Saudi Arabia, the target of the study.

Construction Validity

Although there was a high degree of agreement between the annotators who filtered out tweets not related to health, they did not have health backgrounds. This means that nonhealth tweets may have gotten through this phase. However, this was addressed when the doctors assessed the tweets for health accuracy; they did not identify any tweet as nonhealth related.

This study included the categorization of the authors of tweets into various groups; however, individual health accounts, health organization accounts, and individuals were not externally checked in order to test whether the classification was correct.

In addition, this study intended to examine tweets from Saudi users, specifically during the development of the health lexicon, which was noted in the study's design. However, we cannot guarantee that all tweets had a Saudi origin.

Conclusions

The purpose of this work was to validate the method used to ensure that it was practical for determining the accuracy and the factors associated with the trustworthiness and dissemination of health-related tweets; this was done to provisionally assess factors that may impact on trustworthy tweets and dissemination of tweets. Our results indicate that there may be some clear differences between tweets labeled as trustworthy health information and tweets labeled as untrustworthy health information. They also showed that trusted health professionals were more likely to tweet accurate health information, while exclusively SM-based accounts were more likely to produce untrustworthy tweets. Interestingly, most of the trustworthy tweets were tweeted in the morning, while more of the untrustworthy tweets were tweeted at night. Regarding the dissemination of tweets, there were some features that appeared to be associated with a high dissemination of the tweets. These features appeared at both the tweet-level and user-level analyses.

Due to the limited quantity of data, we cannot have confidence in statistical predictive modelling. The results illustrate that future studies using a large dataset may produce a predictive model for classifying tweets as either trustworthy or untrustworthy.

Acknowledgments

We would like to thank Dr Yousef Albalawi and Dr Khaled Shahat for their valuable advice. In addition, we thank all the doctors who participated in the process of evaluating health tweets in this study. This work was supported, in part, by Taibah University, Al-Ula, Saudi Arabia, and by a grant from the Science Foundation Ireland (grant number: 13/RC/2094).

Conflicts of Interest

None declared.

Multimedia Appendix 1
Complete list of keywords used in the study.
[PDF File (Adobe PDF File), 326 KB - jmir_v21i10e14731_app1.pdf]

Multimedia Appendix 2
Tweet and user features used in the study.
[PDF File (Adobe PDF File), 205 KB - jmir_v21i10e14731_app2.pdf]

References

1. Lowe B, Laffey D. Is Twitter for the birds? J Mark Educ 2011 May 31;33(2):183-192. [doi: 10.1177/0273475311410851]
2. Jiang L, Erdem M. Twitter marketing in multi-unit restaurants: Is it a viable marketing tool? J Foodserv Bus Res 2016 Sep 07;20(5):568-578. [doi: 10.1080/15378020.2016.1222746]
3. Moorhead SA, Hazlett DE, Harrison L, Carroll JK, Irwin A, Hoving C. A new dimension of health care: Systematic review of the uses, benefits, and limitations of social media for health communication. J Med Internet Res 2013 Apr 23;15(4):e85 [FREE Full text] [doi: 10.2196/jmir.1933] [Medline: 23615206]
4. Crook B, Glowacki EM, Suran M, Harris JK, Bernhardt JM. Content analysis of a live CDC Twitter chat during the 2014 Ebola outbreak. Commun Res Rep 2016 Oct 14;33(4):349-355. [doi: 10.1080/08824096.2016.1224171]
5. Wong R, Harris JK, Staub M, Bernhardt JM. Local health departments tweeting about Ebola: Characteristics and messaging. J Public Health Manag Pract 2017;23(2):e16-e24. [doi: 10.1097/PHH.0000000000000342] [Medline: 26334537]
6. McInnes C. 'Add Twitter and stir': The use of Twitter by public authorities in Norway and UK during the 2014-15 Ebola outbreak. Observatorio 2018;12(2):23-46. [doi: 10.15847/obsOBS12220181173]
7. Levac JJ, O'Sullivan T. Social media and its use in health promotion. Interdiscip J Health Sci 2010 Feb 10;1(1):47 [FREE Full text] [doi: 10.18192/riss-ijhs.v1i1.1534]
8. Heldman AB, Schindelar J, Weaver JB. Social media engagement and public health communication: Implications for public health organizations being truly “social”. Public Health Rev 2013 Jun 3;35(1):13. [doi: 10.1007/BF03391698]
9. Omrani AS, Shalhoub S. Middle East respiratory syndrome coronavirus (MERS-CoV): What lessons can we learn? J Hosp Infect 2015 Nov;91(3):188-196. [doi: 10.1016/j.jhin.2015.08.002] [Medline: 26452615]
10. Oyeyemi SO, Gabarron E, Wynn R. Ebola, Twitter, and misinformation: A dangerous combination? BMJ 2014 Oct 14;349:g6178. [doi: 10.1136/bmj.g6178] [Medline: 25315514]
11. Chew C, Eysenbach G. Pandemics in the age of Twitter: Content analysis of tweets during the 2009 H1N1 outbreak. PLoS One 2010 Nov 29;5(11):e14118 [FREE Full text] [doi: 10.1371/journal.pone.0014118] [Medline: 21124761]
12. Alsobayel H. Use of social media for professional development by health care professionals: A cross-sectional Web-based survey. JMIR Med Educ 2016 Sep 12;2(2):e15 [FREE Full text] [doi: 10.2196/mededu.6232] [Medline: 27731855]
13. Omair MA, AlOhaly RY, Alashgar LM. Awareness and misconceptions of female students in King Saud University on systemic lupus erythematosus. Rheumatology (Sunnyvale) 2015;5(3):165 [FREE Full text] [doi: 10.4172/2161-1149.1000165]
14. Jamal A, Khan SA, AlHumud A, Al-Duhyyim A, Alrashed M, Bin Shabr F, et al. Association of online health information-seeking behavior and self-care activities among type 2 diabetic patients in Saudi Arabia. J Med Internet Res 2015 Aug 12;17(8):e196 [FREE Full text] [doi: 10.2196/jmir.4312] [Medline: 26268425]
15. Gold J, Pedrana AE, Sacks-Davis R, Hellard ME, Chang S, Howard S, et al. A systematic examination of the use of online social networking sites for sexual health promotion. BMC Public Health 2011 Jul 21;11:583 [FREE Full text] [doi: 10.1186/1471-2458-11-583] [Medline: 21777470]
16. Capurro D, Cole K, Echavarría MI, Joe J, Neogi T, Turner AM. The use of social networking sites for public health practice and research: A systematic review. J Med Internet Res 2014 Mar 14;16(3):e79 [FREE Full text] [doi: 10.2196/jmir.2679] [Medline: 24642014]
17. Hagg E, Dahinten VS, Currie LM. The emerging use of social media for health-related purposes in low and middle-income countries: A scoping review. Int J Med Inform 2018 Jul;115:92-105. [doi: 10.1016/j.ijmedinf.2018.04.010] [Medline: 29779724]

18. Clement J. Statista. 2019 Aug 09. Leading countries based on number of Twitter users as of July 2019 (in millions) URL: https://www.statista.com/statistics/242606/number-of-active-twitter-users-in-selected-countries/ [accessed 2019-07-18]
19. Worldometers. 2019. Countries in the world by population (2019) URL: https://www.worldometers.info/world-population/population-by-country/ [accessed 2019-07-18]
20. Plotkin Boghardt L. The Washington Institute. Washington, DC: The Washington Institute for Near East Policy; 2013 Dec 09. Saudi Arabia's war on Twitter URL: https://www.washingtoninstitute.org/policy-analysis/view/saudi-arabias-war-on-twitter [accessed 2019-07-19]
21. Mohammed I. The dynamics of the Saudi Twitterverse. In: Mellor N, Rinnawi K, editors. Political Islam and Global Media: The Boundaries of Religious Identity. London, UK: Routledge; 2016:203-218.
22. Alnemer KA, Alhuzaim WM, Alnemer AA, Alharbi BB, Bawazir AS, Barayyan OR, et al. Are health-related tweets evidence based? Review and analysis of health-related tweets on Twitter. J Med Internet Res 2015 Oct 29;17(10):e246 [FREE Full text] [doi: 10.2196/jmir.4898] [Medline: 26515535]
23. Donohue JM, Huskamp HA, Wilson IB, Weissman J. Whom do older adults trust most to provide information about prescription drugs? Am J Geriatr Pharmacother 2009 Apr;7(2):105-116 [FREE Full text] [doi: 10.1016/j.amjopharm.2009.04.005] [Medline: 19447363]
24. Gray NJ, Klein JD, Noyce PR, Sesselberg TS, Cantrill JA. Health information-seeking behaviour in adolescence: The place of the Internet. Soc Sci Med 2005 Apr;60(7):1467-1478. [doi: 10.1016/j.socscimed.2004.08.010] [Medline: 15652680]
25. Hesse BW, Nelson DE, Kreps GL, Croyle RT, Arora NK, Rimer BK, et al. Trust and sources of health information: The impact of the Internet and its implications for health care providers: Findings from the first Health Information National Trends Survey. Arch Intern Med 2005;165(22):2618-2624. [doi: 10.1001/archinte.165.22.2618] [Medline: 16344419]
26. Medlock S, Eslami S, Askari M, Arts DL, Sent D, de Rooij SE, et al. Health information-seeking behavior of seniors who use the Internet: A survey. J Med Internet Res 2015 Jan 08;17(1):e10 [FREE Full text] [doi: 10.2196/jmir.3749] [Medline: 25574815]
27. Worsley A. Perceived reliability of sources of health information. Health Educ Res 1989;4(3):367-376. [doi: 10.1093/her/4.3.367]
28. Kalyanam J, Velupillai S, Doan S, Conway M, Lanckriet G. ACM KDD Workshop on Connected Health in the Big Data Era at the 21st ACM SIGKDD Conference on Knowledge Discovery and Data Mining. 2015 Aug 10. Facts and fabrications about Ebola: A Twitter-based study URL: http://eceweb.ucsd.edu/~gert/papers/KDD_BigCHat_2015.pdf [accessed 2019-09-30]
29. Ghenai A, Mejova Y. Fake cures. Proc ACM Hum Comput Interact 2018 Nov 01;2(CSCW):1-20. [doi: 10.1145/3274327]
30. Yin X, Han J, Yu PS. Truth discovery with multiple conflicting information providers on the Web. IEEE Trans Knowl Data Eng 2008 Jun 12;20(6):796-808. [doi: 10.1109/TKDE.2007.190745]
31. Zhao L, Hua T, Lu C, Chen I. A topic-focused trust model for Twitter. Comput Commun 2016 Feb 15;76:1-11 [FREE Full text] [doi: 10.1016/j.comcom.2015.08.001]
32. Nastasi A, Bryant T, Canner JK, Dredze M, Camp MS, Nagarajan N. Breast cancer screening and social media: A content analysis of evidence use and guideline opinions on Twitter. J Cancer Educ 2018 Jun;33(3):695-702. [doi: 10.1007/s13187-017-1168-9] [Medline: 28097527]
33. Altshuler AL, Gerns Storey HL, Prager SW. Exploring abortion attitudes of US adolescents and young adults using social media. Contraception 2015 Mar;91(3):226-233. [doi: 10.1016/j.contraception.2014.11.009] [Medline: 25537853]
34. Bender JL, Jimenez-Marroquin M, Jadad AR. Seeking support on Facebook: A content analysis of breast cancer groups. J Med Internet Res 2011 Mar 04;13(1):e16 [FREE Full text] [doi: 10.2196/jmir.1560] [Medline: 21371990]
35. Davis MA, Anthony DL, Pauls SD. Seeking and receiving social support on Facebook for surgery. Soc Sci Med 2015 Apr;131:40-47 [FREE Full text] [doi: 10.1016/j.socscimed.2015.02.038] [Medline: 25753284]
36. Guidry JPD, Carlyle KE, LaRose JG, Perrin P, Messner M, Ryan M. Using the health belief model to analyze Instagram posts about Zika for public health communications. Emerg Infect Dis 2019 Jan;25(1):179-180 [FREE Full text] [doi: 10.3201/eid2501.180824] [Medline: 30561302]
37. Schmidt AL, Zollo F, Scala A, Betsch C, Quattrociocchi W. Polarization of the vaccination debate on Facebook. Vaccine 2018 Jun 14;36(25):3606-3612. [doi: 10.1016/j.vaccine.2018.05.040] [Medline: 29773322]
38. Salzmann-Erikson M, Hiçdurmaz D. Use of social media among individuals who suffer from post-traumatic stress: A qualitative analysis of narratives. Qual Health Res 2017 Jan;27(2):285-294. [doi: 10.1177/1049732315627364] [Medline: 26825480]
39. Bhattacharya S, Srinivasan P, Polgreen P. Social media engagement analysis of US federal health agencies on Facebook. BMC Med Inform Decis Mak 2017 Apr 21;17(1):49 [FREE Full text] [doi: 10.1186/s12911-017-0447-z] [Medline: 28431582]
40. Straton N, Hansen K, Mukkamala R, Hussain A, Gronli T, Langberg H. Big social data analytics for public health: Facebook engagement and performance. In: Proceedings of the IEEE 18th International Conference on E-Health Networking, Applications, and Services (Healthcom 2016). New York, NY: IEEE; 2016 Presented at: IEEE 18th International Conference on E-Health Networking, Applications, and Services (Healthcom 2016); September 14-16, 2016; Munich, Germany p. 442-447. [doi: 10.1109/HealthCom.2016.7749497]

41. Kim Y, Huang J, Emery S. Garbage in, Garbage out: Data collection, quality assessment and reporting standards for social media data use in health research, infodemiology and digital disease detection. J Med Internet Res 2016 Mar 26;18(2):e41 [FREE Full text] [doi: 10.2196/jmir.4738] [Medline: 26920122]
42. Abbass MAA, Keshava HB, Delaney CP. The bigger picture: Picking the right soap box-Is it possible to connect with different audience targets (practitioners and patients) from the same platform? Clin Colon Rectal Surg 2017 Sep;30(4):281-290 [FREE Full text] [doi: 10.1055/s-0037-1604258] [Medline: 28924403]
43. Kalyanam J, Mackey TK. A review of digital surveillance methods and approaches to combat prescription drug abuse. Curr Addict Rep 2017 Sep 18;4(4):397-409. [doi: 10.1007/s40429-017-0169-4]
44. Chan WS, Leung AY. Use of social network sites for communication among health professionals: Systematic review. J Med Internet Res 2018 Mar 28;20(3):e117 [FREE Full text] [doi: 10.2196/jmir.8382] [Medline: 29592845]
45. Tang L, Bie B, Park S, Zhi D. Social media and outbreaks of emerging infectious diseases: A systematic review of literature. Am J Infect Control 2018 Sep;46(9):962-972. [doi: 10.1016/j.ajic.2018.02.010] [Medline: 29628293]
46. Fung IC, Duke CH, Finch KC, Snook KR, Tseng P, Hernandez AC, et al. Ebola virus disease and social media: A systematic review. Am J Infect Control 2016 Dec 01;44(12):1660-1671. [doi: 10.1016/j.ajic.2016.05.011] [Medline: 27425009]
47. Charles-Smith LE, Reynolds TL, Cameron MA, Conway M, Lau EH, Olsen JM, et al. Using social media for actionable disease surveillance and outbreak management: A systematic literature review. PLoS One 2015;10(10):e0139701 [FREE Full text] [doi: 10.1371/journal.pone.0139701] [Medline: 26437454]
48. Fu K, Liang H, Saroha N, Tse ZTH, Ip P, Fung IC. How do people react to Zika virus outbreaks on Twitter? A computational content analysis. Am J Infect Control 2016 Dec 01;44(12):1700-1702. [doi: 10.1016/j.ajic.2016.04.253] [Medline: 27566874]
49. Glowacki EM, Lazard AJ, Wilcox GB, Mackert M, Bernhardt JM. Identifying the public's concerns and the Centers for Disease Control and Prevention's reactions during a health crisis: An analysis of a Zika live Twitter chat. Am J Infect Control 2016 Dec 01;44(12):1709-1711. [doi: 10.1016/j.ajic.2016.05.025] [Medline: 27544795]
50. Young R, Tully M, Dalrymple K. #Engagement: use of Twitter chats to construct nominal participatory spaces during health crises. Inform Commun Soc 2017 Mar 27;21(4):499-515. [doi: 10.1080/1369118x.2017.1301518]
51. Miller M, Banerjee T, Muppalla R, Romine W, Sheth A. What are people tweeting about Zika? An exploratory study concerning its symptoms, treatment, transmission, and prevention. JMIR Public Health Surveill 2017 Jun 19;3(2):e38 [FREE Full text] [doi: 10.2196/publichealth.7157] [Medline: 28630032]
52. Juric R, Kim I, Panneerselvam H, Tesanovic I. Analysis of Zika virus tweets: Could Hadoop platform help in global health management. In: Proceedings of the 50th Hawaii International Conference on System Sciences. 2017 Presented at: 50th Hawaii International Conference on System Sciences; January 3-6, 2017; Waikoloa Village, HI. [doi: 10.24251/hicss.2017.395]
53. Park H, Reber BH, Chon M. Tweeting as health communication: Health organizations' use of Twitter for health promotion and public engagement. J Health Commun 2016;21(2):188-198. [doi: 10.1080/10810730.2015.1058435] [Medline: 26716546]
54. Vos SC, Buckner MM. Social media messages in an emerging health crisis: Tweeting bird flu. J Health Commun 2016;21(3):301-308. [doi: 10.1080/10810730.2015.1064495] [Medline: 26192209]
55. Bhattacharya S, Srinivasan P, Polgreen P. Engagement with health agencies on Twitter. PLoS One 2014;9(11):e112235 [FREE Full text] [doi: 10.1371/journal.pone.0112235] [Medline: 25379727]
56. Neiger BL, Thackeray R, Van Wagenen SA, Hanson CL, West JH, Barnes MD, et al. Use of social media in health promotion: Purposes, key performance indicators, and evaluation metrics. Health Promot Pract 2012 Mar;13(2):159-164. [doi: 10.1177/1524839911433467] [Medline: 22382491]
57. Dumbrell D, Steele R. Twitter and health in the Australian context: What types of information are health-related organizations tweeting? In: Proceedings of the 46th Hawaii International Conference on System Sciences. New York, NY: IEEE; 2013 Presented at: 46th Hawaii International Conference on System Sciences; January 7-10, 2013; Maui, HI p. 7-10. [doi: 10.1109/HICSS.2013.578]
58. Ghenai A, Mejova Y. Catching Zika fever: Application of crowdsourcing and machine learning for tracking health misinformation on Twitter. In: Proceedings of the IEEE International Conference on Healthcare Informatics (ICHI). 2017 Presented at: IEEE International Conference on Healthcare Informatics (ICHI); August 23-26, 2017; Park City, UT p. 518-518. [doi: 10.1109/ICHI.2017.58]
59. Passerello GL, Hazelwood JE, Lawrie S. Using Twitter to assess attitudes to schizophrenia and psychosis. BJPsych Bull 2019 Aug;43(4):158-166 [FREE Full text] [doi: 10.1192/bjb.2018.115] [Medline: 30784393]
60. Meadows CZ, Tang L, Liu W. Twitter message types, health beliefs, and vaccine attitudes during the 2015 measles outbreak in California. Am J Infect Control 2019 Jun 29:1. [doi: 10.1016/j.ajic.2019.05.007] [Medline: 31266661]
61. Chung JE. A smoking cessation campaign on Twitter: Understanding the use of Twitter and identifying major players in a health campaign. J Health Commun 2016 May;21(5):517-526. [doi: 10.1080/10810730.2015.1103332] [Medline: 27115179]
62. Stautz K, Bignardi G, Hollands GJ, Marteau TM. Reactions on Twitter to updated alcohol guidelines in the UK: A content analysis. BMJ Open 2017 Feb 28;7(2):e015493 [FREE Full text] [doi: 10.1136/bmjopen-2016-015493] [Medline: 28246145]
63. Broniatowski DA, Jamison AM, Qi S, AlKulaib L, Chen T, Benton A, et al. Weaponized health communication: Twitter bots and Russian trolls amplify the vaccine debate. Am J Public Health 2018 Oct;108(10):1378-1384. [doi: 10.2105/AJPH.2018.304567] [Medline: 30138075]

64. Allem J, Ferrara E, Uppu SP, Cruz TB, Unger JB. E-cigarette surveillance with social media data: Social bots, emerging topics, and trends. JMIR Public Health Surveill 2017 Dec 20;3(4):e98 [FREE Full text] [doi: 10.2196/publichealth.8641] [Medline: 29263018]
65. Jamison AM, Broniatowski DA, Quinn SC. Malicious actors on Twitter: A guide for public health researchers. Am J Public Health 2019 May;109(5):688-692. [doi: 10.2105/AJPH.2019.304969] [Medline: 30896994]
66. Yang K, Varol O, Davis CA, Ferrara E, Flammini A, Menczer F. Arming the public with artificial intelligence to counter social bots. Hum Behav Emerg Technol 2019 Feb 06;1(1):48-61. [doi: 10.1002/hbe2.115]
67. Howard P, Woolley S, Calo R. Algorithms, bots, and political communication in the US 2016 election: The challenge of automated political communication for election law and administration. J Inf Technol Polit 2018 Apr 11;15(2):81-93. [doi: 10.1080/19331681.2018.1448735]
68. Edwards C, Beattie AJ, Edwards A, Spence PR. Differences in perceptions of communication quality between a Twitterbot and human agent for information seeking and learning. Comput Human Behav 2016 Dec;65:666-671. [doi: 10.1016/j.chb.2016.07.003]
69. Albalawi Y, Sixsmith J. Identifying Twitter influencer profiles for health promotion in Saudi Arabia. Health Promot Int 2017 Jun 01;32(3):456-463. [doi: 10.1093/heapro/dav103] [Medline: 26516181]
70. Social Authority. URL: https://followerwonk.com/social-authority [accessed 2019-09-30]
71. PeerIndex. URL: http://peerindex.com [accessed 2019-09-30]
72. Kred. URL: https://home.kred/ [accessed 2019-09-30]
73. Klout. URL: https://khoros.com/platform/klout [accessed 2019-09-30]
74. Suh B, Hong L, Pirolli P, Chi E. Want to be retweeted? Large-scale analytics on factors impacting retweet in Twitter network. In: Proceedings of the IEEE Second International Conference on Social Computing. New York, NY: IEEE; 2010 Presented at: IEEE Second International Conference on Social Computing; August 20-22, 2010; Minneapolis, MN p. 20-22. [doi: 10.1109/SocialCom.2010.33]
75. Kim AE, Hansen HM, Murphy J, Richards AK, Duke J, Allen JA. Methodological considerations in analyzing Twitter data. J Natl Cancer Inst Monogr 2013 Dec;2013(47):140-146. [doi: 10.1093/jncimonographs/lgt026] [Medline: 24395983]
76. Graham M, Hale SA, Gaffney D. Where in the world are you? Geolocation and language identification in Twitter. Prof Geogr 2014 May 19;66(4):568-578. [doi: 10.1080/00330124.2014.907699]
77. Sloan L, Morgan J. Who tweets with their location? Understanding the relationship between demographic characteristics and the use of geoservices and geotagging on Twitter. PLoS One 2015;10(11):e0142209. [doi: 10.1371/journal.pone.0142209] [Medline: 26544601]
78. Twitter Developer. Search tweets URL: https://developer.twitter.com/en/docs/tweets/search/api-reference [accessed 2018-11-15] [WebCite Cache ID 74byAfJy0]
79. Bobicev V, Sokolova M. Thumbs up and down: Sentiment analysis of medical online forums. In: Proceedings of the 2018 Empirical Methods in Natural Language Processing (EMNLP): 3rd Social Media Mining for Health Applications (SMM4H) Workshop & Shared Task. Stroudsburg, PA: Association for Computational Linguistics; 2018 Oct 31 Presented at: 2018 Empirical Methods in Natural Language Processing (EMNLP): 3rd Social Media Mining for Health Applications (SMM4H) Workshop & Shared Task; October 31-November 4, 2018; Brussels, Belgium p. 22-26 URL: https://aclweb.org/anthology/W18-5906 [doi: 10.18653/v1/w18-5906]
80. Cohen J. Weighted kappa: Nominal scale agreement with provision for scaled disagreement or partial credit. Psychol Bull 1968 Oct;70(4):213-220. [doi: 10.1037/h0026256] [Medline: 19673146]
81. The R Core Team. R: A Language and Environment for Statistical Computing. Reference Index. Vienna, Austria: R Foundation for Statistical Computing; 2019 Jul 05. URL: https://cran.r-project.org/doc/manuals/r-release/fullrefman.pdf [accessed 2019-09-30]
82. Castillo C, Mendoza M, Poblete B. Information credibility on Twitter. In: Proceedings of the 20th International Conference on World Wide Web. New York, NY: ACM; 2011 Presented at: 20th International Conference on World Wide Web; March 28-April 1, 2011; Hyderabad, India p. 675-684. [doi: 10.1145/1963405.1963500]
83. Kim A, Miano T, Chew R, Eggers M, Nonnemaker J. Classification of Twitter users who tweet about e-cigarettes. JMIR Public Health Surveill 2017 Sep 26;3(3):e63 [FREE Full text] [doi: 10.2196/publichealth.8060] [Medline: 28951381]
84. Morris M, Counts S, Roseway A, Hoff A, Schwarz J. Tweeting is believing?: Understanding microblog credibility perceptions. In: Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work. New York, NY: ACM; 2012 Presented at: ACM 2012 Conference on Computer Supported Cooperative Work; February 11-15, 2012; Seattle, WA p. 441-450. [doi: 10.1145/2145204.2145274]
85. Zhuang K, Shen H, Zhang H. User spread influence measurement in microblog. Multimed Tools Appl 2016 Aug 4;76(3):3169-3185. [doi: 10.1007/s11042-016-3818-z]
86. Riquelme F, González-Cantergiani P. Measuring user influence on Twitter: A survey. Inf Process Manag 2016 Sep;52(5):949-975. [doi: 10.1016/J.IPM.2016.04.003]
87. El Gayar N, Suen CY, editors. Computational Linguistics, Speech and Image Processing for Arabic Language. Singapore: World Scientific; 2017.

88. Al-Nazer A, Gutub A. Exploit kashida adding to Arabic e-text for high-capacity steganography. In: Proceedings of the Third International Conference on Network and System Security. New York, NY: IEEE; 2009 Presented at: Third International Conference on Network and System Security; October 19-21, 2009; Gold Coast, Australia p. 447-451. [doi: 10.1109/NSS.2009.21]
89. Volkova S, Van Durme B. Online Bayesian models for personal analytics in social media. In: Proceedings of the Twenty-Ninth Association for the Advancement of Artificial Intelligence (AAAI) Conference on Artificial Intelligence. Palo Alto, CA: AAAI Press; 2015 Presented at: Twenty-Ninth Association for the Advancement of Artificial Intelligence (AAAI) Conference on Artificial Intelligence; January 25-30, 2015; Austin, TX p. 2325-2331 URL: http://www.cs.jhu.edu/~vandurme/papers/online-bayesian-models-for-personal-analytics-in-social-media.pdf
90. Schwartz HA, Eichstaedt JC, Kern ML, Dziurzynski L, Ramones SM, Agrawal M, et al. Personality, gender, and age in the language of social media: The open-vocabulary approach. PLoS One 2013;8(9):e73791 [FREE Full text] [doi: 10.1371/journal.pone.0073791] [Medline: 24086296]
91. An L, Yi X, Han Y, Li G. An influence prediction model for microblog entries on public health emergencies. Data Inf Manag 2019;3(2):102-115. [doi: 10.2478/dim-2018-0013]
92. Prabhu V, Lee T, Loeb S, Holmes JH, Gold HT, Lepor H, et al. Twitter response to the United States Preventive Services Task Force recommendations against screening with prostate-specific antigen. BJU Int 2015 Jul;116(1):65-71 [FREE Full text] [doi: 10.1111/bju.12748] [Medline: 24661474]
93. Lancaster GA, Dodd S, Williamson PR. Design and analysis of pilot studies: Recommendations for good practice. J Eval Clin Pract 2004 May;10(2):307-312. [doi: 10.1111/j.2002.384.doc.x] [Medline: 15189396]
94. de Winter JCF, Dodou D. Five-point Likert items: t test versus Mann-Whitney-Wilcoxon. Pract Assess Res Eval 2010;15(11):1-12 [FREE Full text]
95. Mehta CR, Patel NR. Algorithm 643: FEXACT: A Fortran subroutine for Fisher's exact test on unordered r×c contingency tables. ACM Trans Math Softw 1986 Jun;12(2):154-161. [doi: 10.1145/6497.214326]
96. Clarkson DB, Fan Y, Joe H. A remark on algorithm 643: FEXACT: An algorithm for performing Fisher's exact test in r x c contingency tables. ACM Trans Math Softw 1993;19(4):484-488. [doi: 10.1145/168173.168412]
97. Yoo E, Rand W, Eftekhar M, Rabinovich E. Evaluating information diffusion speed and its determinants in social media networks during humanitarian crises. J Oper Manag 2016 Jul 09;45(1):123-133. [doi: 10.1016/j.jom.2016.05.007]

Abbreviations

API: application programming interface
CDC: US Centers for Disease Control and Prevention
F1: number of followers
F3: number of followees
FT1: number of tweets favorited by the author
FT2: number of times that author's tweets are favorited
FT6: number of original tweets posted by the author that are favorited
LDH: local health department
M1: number of tweets where the author mentioned other users
M2: number of unique users mentioned by the author
MERS: Middle East respiratory syndrome
MH5: unique keyword count in hashtags set in original tweets posted by the author
MWW: Mann-Whitney-Wilcoxon
OT3: number of hashtags in the author's tweets
RM1: number of retweeted tweets by the author where the user mentioned other users
RMH5: unique hashtag count in tweets that were retweeted by the author
RMM5: unique mentions in retweeted tweets by the author
RP1: number of reply-to tweets posted by the author
RT1: number of tweets that the author retweeted
RT2: number of author's tweets retweeted by other users
SM: social media
SSI: ratio of original tweets posted by the author to tweets retweeted by the author
TFF: ratio of followers to followees

Edited by G Eysenbach; submitted 19.05.19; peer-reviewed by JP Allem, K Reuter; comments to author 18.06.19; revised version received 02.09.19; accepted 03.09.19; published 08.10.19.

Please cite as:
Albalawi Y, Nikolov NS, Buckley J
Trustworthy Health-Related Tweets on Social Media in Saudi Arabia: Tweet Metadata Analysis
J Med Internet Res 2019;21(10):e14731
URL: https://www.jmir.org/2019/10/e14731
doi:10.2196/14731
PMID:31596242

Review

Electronic Patient-Generated Health Data to Facilitate Disease Prevention and Health Promotion: Scoping Review

Vasileios Nittas[1], BSc, MSc; Penny Lun[2], BA, MA; Frederic Ehrler[3], BSc, MSc, PhD; Milo Alan Puhan[1], Prof Dr, MD; Margot Mütsch[1], MPH, MSc, PhD

[1]Epidemiology, Biostatistics and Prevention Institute, University of Zurich, Zurich, Switzerland

[2]Geriatric Education and Research Institute, Singapore, Singapore

[3]Division of Medical Information Sciences, University Hospitals of Geneva, Geneva, Switzerland

Corresponding Author:
Vasileios Nittas, BSc, MSc
Epidemiology, Biostatistics and Prevention Institute
University of Zurich
Hirschengraben 84
Zurich, CH-8001
Switzerland
Phone: 41 446344946
Email: vasileios.nittas@uzh.ch

Abstract

Background: Digital innovations continue to shape health and health care. As technology socially integrates into daily living, the lives of health care consumers are transformed into a key source of health information, commonly referred to as patient-generated health data (PGHD). With chronic disease prevalence signaling the need for a refocus on primary prevention, electronic PGHD might be essential in strengthening proactive and person-centered health care.

Objective: This study aimed to review and synthesize the existing literature on the utilization and implications of electronic PGHD for primary disease prevention and health promotion purposes.

Methods: Guided by a well-accepted methodological framework for scoping studies, we screened MEDLINE, CINAHL, PsycINFO, Scopus, Web of Science, EMBASE, and IEEE Digital Library. We hand-searched 5 electronic journals and 4 gray literature sources, additionally conducted Web searches, reviewed relevant Web pages, manually screened reference lists, and consulted authors. Screening was based on predefined eligibility criteria. Data extraction and synthesis were guided by an adapted PGHD-flow framework. Beyond initial quantitative synthesis, we reported narratively, following an iterative thematic approach. Raw data were coded, thematically clustered, and mapped, allowing for the identification of patterns.

Results: Of 183 eligible studies, targeting knowledge and self-awareness, behavior change, healthy environments, and remote monitoring, most literature (125/183, 68.3%) addressed weight reduction, either through physical activity or nutrition, applying a range of electronic tools from socially integrated to full medical devices. Participants generated their data actively (100/183, 54.6%), in combination with passive sensor-based trackers (63/183, 34.4%) or entirely passively (20/183, 10.9%). The proportions of active and passive data generation varied strongly across prevention areas. Most studies (172/183, 93.9%) combined electronic PGHD with reflective, process guiding, motivational and educational elements, highlighting the role of PGHD in multicomponent digital prevention approaches. Most of these interventions (110/183, 60.1%) were fully automatized, underlining broader trends toward low-resource and efficiency-driven care. Only a fraction (47/183, 25.6%) of studies provided indications on the impact of PGHD on prevention-relevant outcomes, suggesting overall positive trends, especially on vitals (eg, blood pressure) and body composition measures (eg, body mass index). In contrast, the impact of PGHD on health equity remained largely unexplored. Finally, our analysis identified a list of barriers and facilitators clustered around data collection and use, technical and design considerations, ethics, user characteristics, and intervention context and content, aiming to guide future PGHD research.

Conclusions: The large, heterogeneous volume of the PGHD literature underlines the topic's emerging nature. Utilizing electronic PGHD to prevent diseases and promote health is a complex matter owing to mostly being integrated within automatized and multicomponent interventions. This underlines trends toward stronger digitalization and weaker provider involvement. A PGHD use that is sensitive to identified barriers, facilitators, consumer roles, and equity considerations is needed to ensure effectiveness.

(J Med Internet Res 2019;21(10):e13320) doi:10.2196/13320

KEYWORDS

patient-generated health data; personal health information; consumer health information; primary prevention; health promotion; telemedicine; mobile health; medical informatics; eHealth

Introduction

Background

The emergence of digital health innovations is expected to continue shaping the organization and delivery of health services [1,2]. As technology integrates into multiple domains of daily living, its potential for disrupting health systems and societal impact rapidly expands [1,3]. On an individual level, the uptake of smart and wearable technology pushes the boundaries of self-quantification and generates novel opportunities to monitor and promote health [2,4]. These developments gradually transform the lives of health care consumers into key health information sources. The output is commonly referred to as patient-generated health data (PGHD) [5].

Electronic and Patient-Generated Health Data

A landmark whitepaper by the US Office of the National Coordinator for Health Information Technology defines PGHD as health-related information created by patients or their designees outside traditional health care contexts [6]. The current emergence of PGHD can be partially attributed to 2 dominant digitalization trends: the societal integration of mobile phones and the growing health-related use of Web-based media [7-9]. Preinstalled mobile phone applications and integrated sensors enable the continuous measurement of physical, mental, social, and environmental health parameters, whereas online platforms and social media increasingly become a place for health communication and depositories of large data volumes [7,8,10]. These trends gradually transform consumers from passive recipients to active agents of their health [1]. Acknowledging the wide social integration of mobile devices, the generation of one's own health information might be a potential way to engage medically underserved populations and close long-lasting inequity gaps [11].

Digital and Proactive Prevention

As the prevalence of chronic diseases continues to rise, many health care systems face unprecedented challenges that deem it necessary to refocus on prevention [12]. Without disregarding the challenges of electronic PGHD, generating one's own health information might provide an incentive for behavioral change, facilitating health literacy and knowledge exchange [13,14]. Data sharing can in turn trigger personalized feedback, customized health plans, and tailored persuasive health promotion techniques [14,15]. In other words, PGHD can contribute to proactive, informed, and prevention-focused health systems, as well as personalized and collaborative care [16-18]. Despite those benefits, systematically and comprehensively synthesized knowledge on the use of such data for primary disease prevention and health promotion purposes seems to be lacking.

Objectives

Our overarching objective targets the synthesis of the literature on the overall utilization of electronic PGHD for primary disease prevention and health promotion purposes. Specific objectives include (1) providing an overview of applied PGHD types and tools, as well as their aims, purposes, and contexts, (2) exploring health care consumer, provider, and technology responsibilities, as well as potential interactions among them, and (3) synthesizing broader implications of electronic PGHD on health outcomes and equity.

Methods

Methodological Framework

Overview

Our methodology was guided by Arksey and O'Malley's framework for scoping studies and Levac, Colquhoun, and O'Brien's conceptual extensions [19,20]. We divided our approach accordingly into 6 steps, described separately in the following sections. Study quality and evidence strength assessment fall beyond the aims of a scoping review and were not performed [19]. A detailed description of our methodological and conceptual background has been published elsewhere [21]. Protocol deviations and their justifications are provided in Multimedia Appendix 1.

Step 1: Identifying the Research Question

Our research question was formed by an iterative process by getting acquainted with the literature, identifying existing evidence gaps, as well as by regular exchange and expert consultation. Our question consisted of the 3 previously mentioned study objectives across the underlying dimensions of (1) data generation and collection, (2) sharing or communication, (3) interpretation, and (4) utilization. We narrowed the definition of electronic PGHD to data generated by consumer-facing means, excluding information that was collected through standardized, provider-driven methods, such as predefined questionnaires [22], which is justified on the very nature of primary disease prevention and health promotion that ideally requires an active patient. Similarly, we used the term health care consumer, instead of patient, as our target population consists of individuals free of any International Classification of Disease–coded conditions. Nonetheless, we kept the *patient* in PGHD, as that is an already coined term. The term provider is conceptualized as any professional who is responsible for offering health and well-being–related services.

Step 2: Identifying Relevant Studies

With the support of a specialized librarian and preliminary literature review, we developed an extensive and purposively sensitive search strategy, applied to 7 electronic databases that included MEDLINE, Cumulative Index to Nursing and Allied

Health Literature, PsycInfo, Scopus, Web of Science, EMBASE, and Institute of Electrical and Electronics Engineers Digital Library. The searches were conducted on February 1, 2018. We additionally hand-searched 5 key electronic journals and 4 gray literature sources, complemented by Web searches, using the first 10 page results of 3 engines and thorough screenings of 6 relevant Web pages. Our last research steps consisted of (1) the manual reference list screening of all eligible studies and (2) author consultations, requesting input on potentially missed or unpublished work. A more detailed description of our study identification strategy is provided in previously published protocol [21].

The full search strategy and search terms are provided in Multimedia Appendix 2.

Step 3: Study Selection

A total of 2 members of the research team (VN and PL) independently conducted a screening of the titles and abstracts, as well as full text screening against a set of predefined eligibility criteria (Textbox 1). Not fulfilling all conditions below led to exclusion.

Textbox 1. Study eligibility criteria.

- Addresses electronic patient-generated health data (PGHD), as defined by this review, and additionally does the following:
  - Includes at least one sentence on the electronic PGHD tool or type.
  - Includes at least one sentence on how these are used or created.
  - Addresses PGHD that are available in a digital format at the point of utilization for intended health-related purposes, irrespective of the generation process.
- Has a main focus on primary prevention and health promotion and falls within one the following domains:
  - Preventing initial occurrence of disease in healthy or high-risk individuals.
  - Mitigating risk in healthy or high-risk individuals.
  - Promoting existing health.
- Describes, explores, and analyzes some form of health care consumer and provider involvement, where *provider* can be a technology or a human provider.
- Addresses an adult population.
- Is a primary study published in English or German between January 1, 2003, and January 31, 2018.

Step 4: Charting the Data

Data extraction was conducted by 2 reviewers (VN and PL), guided by a predefined, flexible data extraction form, to capture the review's objectives and corresponding research questions. The final form was refined and validated through consultation and expert feedback. Impact data were broadly extracted in terms of significance and direction. Equity data were extracted according to Cochrane Equity Group recommendations [23]. The 2 reviewers (VN and PL) initially tested the form on a random sample of 5 studies, following immediate comparisons and adjustments [20,24]. Owing to the large volume of the included literature, extractions were divided among the 2 reviewers. To minimize subjective bias, 27.3% (50/183) of all completed data extraction forms were randomly selected for validation by a third reviewer who added comments and amendments. Recommended changes were discussed and integrated with consensus.

Step 5: Collating, Summarizing, and Reporting the Results

The whole process, including data charting (Step 4), was guided by an adapted PGHD flow framework, provided in Multimedia Appendix 3 [6,21]. Our adapted version visualizes the flow of PGHD from the consumer, passing through intermediaries (technology or health care provider) and back to the consumer in the form of prevention and health promotion impact. Initial synthesis was quantitative, aiming to provide a descriptive summary of study and participant characteristics, as well as the extent, scope, and nature of the existing literature. Further synthesis was qualitative and followed an iterative thematic approach [20]. Raw data were coded, thematically clustered, and transformed into conceptual maps that facilitated the identification of patterns. The entire process, including screening (Step 3) and data extraction (Step 4), was conducted in Covidence (Cochrane), a Web-based systematic reviewing tool, and DistillerSR (Evidence Partners), a multi-functional software for all types of literature reviews. Reporting is based on the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement [25].

Step 6: Consultation

An external PGHD expert was consulted twice during the conceptualization stage who provided content-related feedback. A total of 3 stakeholders, a provider-partner (physician) and 2 consumer-partners, were consulted during the final manuscript preparation stages to ensure that our interpretations were relevant and understandable.

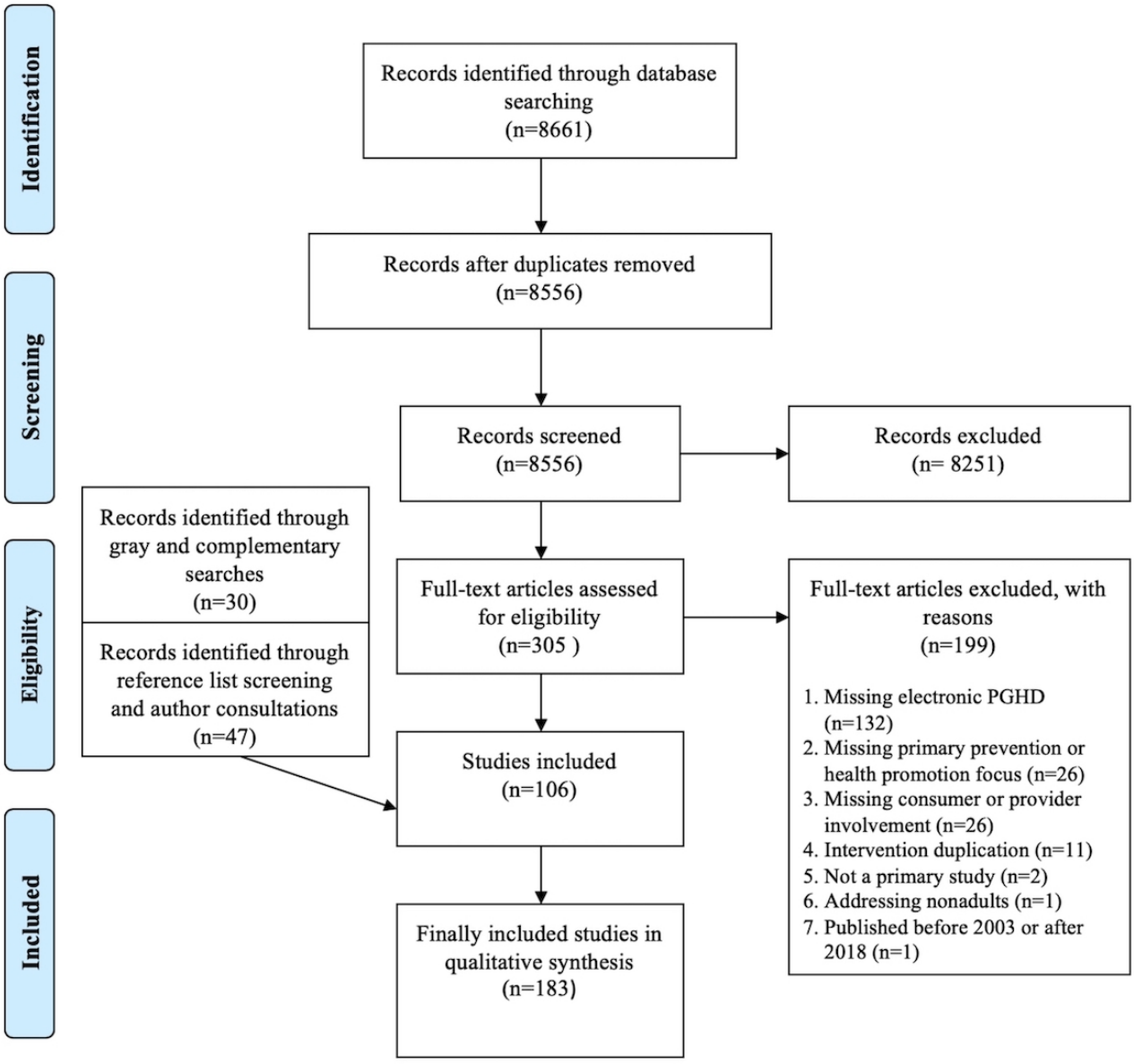
Results

Summary

The deduplicated database search resulted in 8556 citations, which were screened by titles and abstracts. Full-text appraisal

was deemed eligible for 305 studies, of which 199 did not fulfill our inclusion criteria. In total, the electronic database searches yielded 106 included studies. Interrater agreement reached 84% (42/50) (k=0.411) for a sample of 50 studies during title and abstract screening and 93% (14/15) (k=0.636) for a sample of 15 studies during full-text review. Complementary searches, including hand searching and searching gray literature sources, led to the inclusion of 30 studies, whereas reference list screenings and author consultations led to 47 additional studies, resulting in a total of 183 inclusions. A list of all excluded references at full-text screening, including justifications, is provided in Multimedia Appendix 4. A list of included studies and their extracted study characteristics are provided in Multimedia Appendices 5 and 6, respectively. The PRISMA flow chart [25] in Figure 1 summarizes the whole study process.

Figure 1. Preferred Reporting Items for Systematic Reviews and Meta-Analyses flow chart. PGHD: patient-generated health data.



Study Characteristics

Most eligible studies were published as scientific journal articles (n=162), followed by sections of conference proceeding collections (n=13) and published theses (n=8). With an average of 22 studies per year, most were published from 2011 onward, whereas the number of publications averaged to around 6 studies a year between 2003 and 2010. More than half of the studies were conducted in North America (n=107), with 105 from the United States and 2 from Canada. European research followed with 38 studies, most of which were conducted in the United Kingdom. The remaining were conducted in Australia and New Zealand (n=18), Asia (n=13), and the Middle East (n=1). A total of 6 studies had an international scope. Randomized controlled trials constituted most of the applied methodologies (n=93), followed by quantitative nonrandomized approaches (n=47), mixed-method designs (n=30), and purely qualitative methodologies (n=13). The majority aimed to demonstrate effectiveness and efficacy (n=99), followed by mixed and purely exploratory aims (n=52), whereas less than a quarter explored feasibility and acceptability of interventions (n=32). The duration of identified studies ranged from a single examination

to up to 2 years. More detailed information on study characteristics, including percentages, is provided in Multimedia Appendix 5. Participant characteristics are summarized in Multimedia Appendix 7.

Electronic Patient-Generated Health Data—Enabled Prevention Areas and Aims

The most commonly addressed prevention area was weight management, primarily in the form of physical activity and nutrition, which consists of 68.3% (125/183) of the identified literature. This is followed by 12.0% (22/183) of studies with a broader focus on health and well-being. These studies did not exclusively focus on one prevention area and included combinations of chronic and infectious disease, as well as mental health. About 7.7% (14/183) of the literature addressed cardiometabolic health, whereas 7.1% (13/183) focused on substance use prevention, targeting tobacco and alcohol. Healthy aging, such as prevention of falls, cognitive decline, and bone health, was the subject of 6 studies (6/183, 3.3%), followed by 2 studies on breastfeeding (2/183, 1.1%) and 1 study on skin cancer prevention (1/183, 0.5%). Multimedia Appendix 8 provides a list of included studies grouped by prevention areas.

We continued our analysis by synthesizing information on the aims of generating and sharing PGHD for primary disease prevention and health promotion purposes. We identified that enabling health consumers to generate their own health information aims at (1) promoting healthy behavior (142/183, 77.5%), (2) increasing health knowledge and self-awareness (120/183, 65.5%), (3) enabling healthy environments (60/183, 32.7%), and (4) enhancing remote monitoring (20/183, 10.9%). Most studies (134/183, 73.2%) targeted 2 or more of those aims. A similar pattern was observed within prevention areas, with health behavior change and knowledge gain being the most commonly addressed aims. This was not the case for substance use prevention, where enabling healthy environments outweighed knowledge gain. Not all studies adhered strictly to those aims, with 21.3% (39/183) deviating from purely preventive purposes and additionally using PGHD as outcome measures, for example, to quantify the effects of interventions and for secondary analyses. Table 1 describes the 4 identified aims and provides examples from the literature.

Table 1. Description and examples of patient-generated health data aims.

Aim	Description	Example from the literature
Increase health knowledge and self-awareness	Increase in knowledge and cognition about one's health, well-being, and behavior, with no particular focus on how to translate this knowledge into action and concrete behavior	Participants record dietary intake and receive weekly feedback with summaries on their fruit, vegetable, and junk food intake [26]
Promote healthy behavior	Help translate one's knowledge into action, behavior change, and skill development, targeting health improvement and maintenance	Participants record dietary intake and receive nutritional feedback and additional individual dietary targets, recipes, and a meal plan for achieving those [27]
Enable healthy environments	Enable environments and contexts that facilitate health and well-being	Participants record physical activity in a digital partnership with family members or friends, creating an environment of healthy social pressure and support [28]
Enhance remote monitoring	Enable the remote monitoring of individual health and well-being parameters, by health care and wellness providers	Participants record blood pressure, blood glucose, weight, and body fat at home and sent data electronically to medical professional who monitors and provides personalized physical activity plans [29]

The Role of Consumers

Successful prevention undoubtedly requires a clear definition of health care consumer responsibilities. Our analysis identified 3 broad consumer roles. The first consisted of passive PGHD generation (20/183, 10.9%), in which consumers used sensor-based devices to automatically collect and transmit information. Such an approach was predominantly applied in physical activity, weight loss, and overall health and well-being, capturing data that did not require manual entries, such as step counts, heart rate, and sleep quality. The more common second and third roles consisted of fully (100/183, 54.6%) or partially (63/183, 34.4%) active consumers, requiring occasional to regular actions. Active consumer involvement is key for capturing data that are not easily captured automatically, such as consumed meals and the quantity of smoked cigarettes. The term *partially active* describes all approaches that involve both active and passive data generation. That includes anything that is not exclusively sensor-based, nor exclusively dependent on manual entries. Partially active data generation was highly prevalent in prevention areas that often require the simultaneous collection of multiple heterogenous measures (eg, steps, food intake, blood pressure and blood glucose), which was the case in weight loss, physical activity, nutrition, and cardiometabolic health. Table 2 summarizes consumer roles and provides illustrative literature examples. Table 3 provides the spread of consumer roles across identified prevention areas.

Table 2. Patient-generated health data–related consumer roles and examples.

Consumer roles	Examples from the literature
Data generation roles	
Fully active data generation	Take picture of meal and optionally add descriptions, visit website to add further contextual information [30]
Partially active data generation	Manually record stress levels and automatically capture data by wearing heart monitor [31]
Passive data generation	Carry mobile phone an physical activity monitor that generates PGHD[a] automatically [32]
Data sharing roles	
Low-intensity data sharing	PGHD automatically stored on mobile phone based database and automatically transmitted in an encrypted manner [33]
High-intensity data sharing	Share data manually from monitors to website (directly or via a docking station) [34-35]

[a]PGHD: patient-generated health data.

Data generation is often followed by data sharing to third parties or across devices and storage locations. Information on data sharing was provided by about 73.2% (134/183) of the literature. We defined high-intensity sharing as any transmission of PGHD that requires concrete consumer action. High-intensity sharing was applied in 91 studies (91/183, 49.7%). Half of those (39/91, 43%) indicated more demanding actions requiring the active transfer of PGHD to external devices (eg, external computer) or storage locations (eg, server and website). In contrast, low-intensity sharing describes the automatic transmission of PGHD, which was applied in 43 (43/183, 23.5%) studies. We did not identify any difference of distribution between higher or lower sharing intensity across most prevention areas, except for cardiometabolic health and weight loss (Table 3). In cardiometabolic health, most studies described high-intensity data sharing. This can be attributed to the frequent use of less connected devices, such as blood pressure monitors and glucometers. An opposite trend was observed in weight loss interventions that tended to adopt low-intensity sharing, which could be attributable to the sophistication and good interoperability of fitness trackers. Among the studies (n=163/183) with fully or partially active consumers, 137 provided clear information on the frequency of PGHD sharing. Out of those, 110 (110/137, 80.3%) described daily and 27 (27/137, 19.7%) less than daily sharing frequency.

On the basis of the initial framework by Shapiro et al, we developed an extended and more comprehensive conceptual framework [6,21]. With consumer roles as a starting point, our framework visualizes the generation and flow of electronic PGHD and is adapted to the context of primary disease prevention and health promotion. The framework, illustrated in Figure 2, organizes the study's key results and visualizes identified patterns.

Our enriched framework shows that the 3 identified consumer roles are linked to different PGHD tools, ultimately creating different clusters of data. Although we identified interventions stopping at that level (single-component), the majority entailed additional intervention components (multicomponent), with and without human provider involvement. Thus, prevention and health promotion impact can be achieved at 3 levels, as the lower arrows indicate. Across the different elements, from the consumer to the provider, 5 common areas of barriers and facilitators can inhibit or promote the effective use of electronic PGHD. The framework additionally visualizes the link between identified PGHD aims and additional intervention components, as well as the involvement of health care providers. This framework fulfills the function of providing a simplified process overview, ultimately fostering a better understanding of PGHD utilization across prevention areas. All framework components are detailed throughout the results section.

Table 3. Distribution overview of key themes across prevention areas.

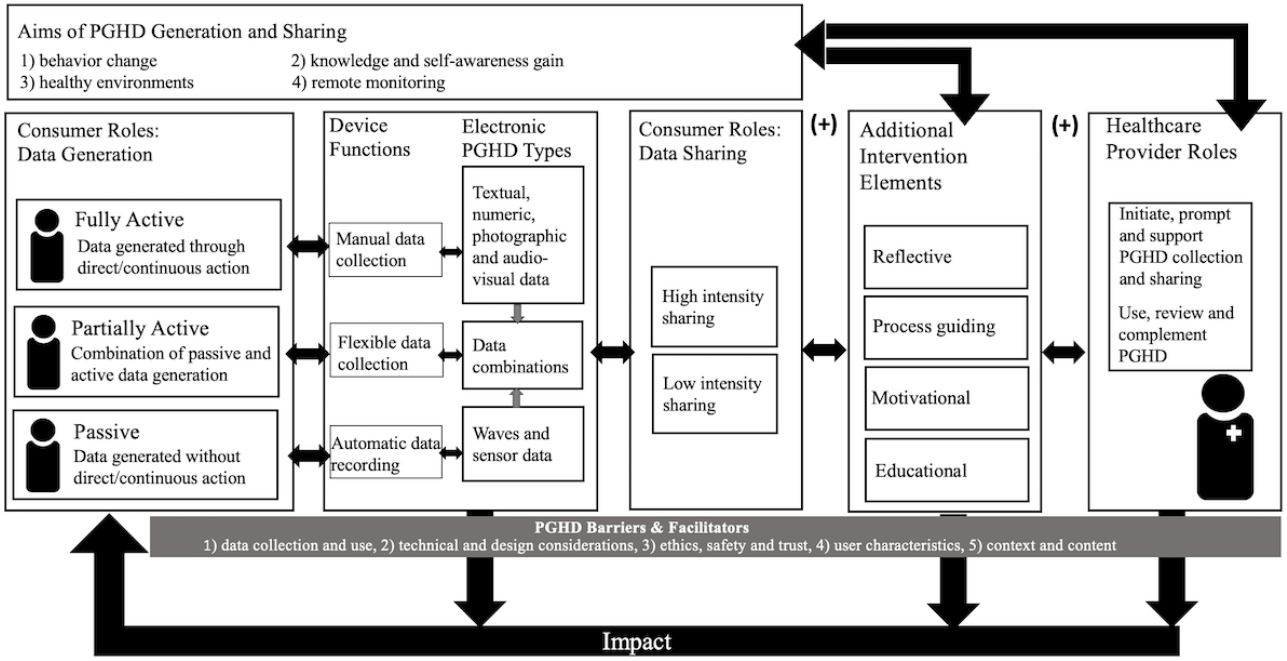
Prevention areas	Weight control, physical activity, nutrition (n=125), n (%)	Overall health and well-being (n=21), n (%)	Cardio-metabolic health (n=14), n (%)	Substance use (smoking and alcohol; n=14), n (%)	Healthy aging (n=6), n (%)	Breastfeeding (n=2), n (%)	Skin cancer (n=1), n (%)
Consumer roles (data generation)[a]							
Active data generation	65 (52.0)	11 (52)	4 (29)	14 (100)	3 (50)	2 (100)	1 (100)
Partially active data generation	46 (36.8)	7 (33)	9 (64)	0 (0)	1 (17)	0 (0)	0 (0)
Passive data generation	14 (11.2)	3 (15)	1 (7)	0 (0)	2 (33)	0 (0)	0 (0)
Consumer roles (data sharing)[a]							
High-intensity data sharing	43 (34.4)	9 (43)	8 (58)	4 (29)	1 (17)	0 (0)	1 (100)
Low-intensity data sharing	51 (40.8)	8 (38)	3 (21)	4 (29)	2 (33)	0 (0)	0 (0)
Unclear or not described	31 (24.8)	4 (19)	3 (21)	6 (42)	3 (50)	2 (100)	0 (0)
Health care provider roles							
Support and motivate PGHD[b]	2 (1.6)	3 (14)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Review and analyze PGHD	23 (18.4)	1 (5)	5 (36)	3 (21)	1 (17)	1 (50)	0 (0)
Support and Motivate PGHD and review and analyze PGHD combined	12 (9.6)	0 (0)	4 (29)	1 (7)	0 (0)	0 (0)	0 (0)
Non-PGHD-related involvement	14 (11.2)	2 (10)	1 (7)	0 (0)	0 (0)	0 (0)	0 (0)
No involvement at all	74 (59.2)	15 (71)	4 (28)	10 (72)	5 (83)	1 (50)	1 (100)
Utilized hardware[c]							
Nonhealth-related (eg, phone)	104 (83.2)	18 (86)	12 (86)	14 (100)	5 (83)	2 (100)	1 (100)
Health-related (eg, pedometer)	65 (52.0)	12 (87)	10 (71)	0 (0)	2 (33)	0 (0)	0 (0)
Medical (eg, glucometer)	2 (1.6)	3 (14)	5 (36)	5 (36)	0 (0)	0 (0)	0 (0)
Additional intervention components[a,c]							
Reflective	113 (90.4)	16 (76)	13 (93)	11 (79)	5 (83)	2 (100)	0 (0)
Process guiding	99 (79.2)	9 (43)	13 (93)	12 (86)	5 (83)	1 (50)	1 (100)
Motivational	88 (70.4)	8 (38)	7 (50)	13 (93)	4 (67)	0 (0)	1 (100)
Educational	84 (67.2)	6 (29)	10 (71)	10 (71)	2 (33)	2 (100)	0 (0)

[a]Consumer roles are described in patient-generated health data–related consumer roles and examples table. Additional intervention components are defined in detail in descriptions of intervention components.

[b]PGHD: patient-generated health data.

[c]Studies were assigned to multiple hardware and additional intervention component categories, for which rows do not add up to 100%.

Figure 2. Enriched conceptual framework of electronic patient-generated health data (PGHD) flow and use for primary disease prevention and health promotion.



Electronic Patient-Generated Health Data Tools, Their Functions, and Data Types

Prevention-targeted PGHD were generated through 3 broad types of hardware, often used in combination. The first included nonhealth-related products that are mostly well integrated into daily living (157/183, 85.8%), such as computers and mobile phones. The second entailed health-related devices that are less societally integrated (90/183, 49.2%), such as pedometers and heart rate monitors. The third included more specialized medical devices (15/183, 8.2%), such as glucometers and blood pressure monitors. Beyond manual and automatic PGHD collection, their most common functions included the provision of additional intervention components, such as goal setting and reminders, analysis and visualization of data, provision of feedback, sharing and storage of PGHD, and communication and interaction with third parties. Another key function was the provision of cues and visualizations, such as using color schemes, pictures, avatars, and other virtual elements to support the interpretation of PGHD. Table 3 provides the distribution of utilized hardware across prevention areas. Multimedia Appendix 9 offers more detailed information on identified tools and their functions. Most studies (175/183, 95.6%) have additionally described the use of software, such as apps, mobile- and Web-based programs, Web-based platforms and websites, social media and forums, device-installed software, and email and text messaging. The remaining 4.4% (8/183) did not explicitly mention any utilized software.

The identified electronic PGHD were categorized in 4 broad types. Most studies (78/183, 42.6%) addressed textual or numerical data, requiring manual entry and an active consumer. This was followed by waves or signals (22/183, 12.0%) that did not require manual collection and audiovisual (video) (4/183, 2.2%) as well as photographic (2/183, 1.1%) PGHD, which again required an active user. Textual or numerical data were mostly used in weight control, substance use prevention, and healthy aging. Photographic data were applied in healthy eating, whereas audiovisual PGHD were commonly utilized in smoking and alcohol prevention. Waves and signals were primarily applied in weight control, well-being, cardiometabolic health, and healthy aging. Finally, almost half of the studies used 2 or more forms of digital PGHD (77/183, 42.1%), with the most common combinations being that of textual or numerical with waves or signals (58/77, 75%) and textual or numerical with photographic data (7/77, 9%). Textual or numerical with waves or signals was used across the spectrum of health domains from weight control to diabetes prevention and usually included initial sensor data that were then manually recorded by users. Textual or numerical with photographic data was applied in dietary interventions, where users took pictures of their meals and added descriptions.

Electronic Patient-Generated Health Data Use: Additional Intervention Components

In 172 (172/183, 93.9%) of the identified studies, PGHD were embedded in larger multicomponent preventive interventions. Our analysis identified 4 overarching components to which PGHD were combined with, categorized as (1) reflective, (2) process guiding, (3) motivational, and (4) educational. Their descriptions and examples are provided in Textbox 2. Most interventions entailed at least one (162/183, 88.5%) purely reflective component, whereas 77.6% (142/183) included at least one process guiding component. Motivational components were included in 67.2% (123/183) of all interventions, with social support (eg, online peer interaction) constituting 33.3% (61/183) of those. Finally, educational components were identified in 63.4% (116/183) of all studies. A distinct element of educational components was the provision of support for making sense of one's health data, through *a-priori* training or instructions, as well as targeted immediate or retrospective feedback. Table 3 provides the distribution of additional

intervention components across prevention areas. The boundaries between those 4 components were not fixed, and each study was often assigned to more than 1 category. Most of these components (147/172, 85.5%) were entirely or partially tailored to individual participants, whereas a relatively smaller proportion (25/172, 14.5%) were predominantly nontailored or unclear. The overlap between identified aims for collecting PGHD and additional interventions components (eg, educational components and the aim of knowledge enhancement, motivational components, and aim of behavior change) suggested that underlying reasons for generating and using PGHD for preventive purposes might influence and be influenced by the availability of these additional intervention components. Additional intervention components are defined in detail in Textbox 2.

Textbox 2. Descriptions of intervention components.

1. Reflective: All intervention components that are based on simple feedback of generated patient-generated health data (PGHD), with no additional educational information on how these are to be interpreted and applied. Examples include PGHD reports and summaries, as well as access to unstructured data.
2. Process guiding: All intervention components that aim to provide general support on the generation of PGHD, the use of technology, the compliance to intervention guidelines, and the response to problems arising from these processes. They include technical advice, instructions on when and how to collect and share, and problem-solving advice. Guidance in understanding and applying PGHD falls out this category's scope (see point 4. Educational).
3. Motivational: All intervention components that are based on techniques that target the motivation of users to collect PGHD and apply those for healthy behavior changes. They include the provision of rewards and incentives, persuasion techniques, goal setting, reminders, motivational feedback, social support, as well as entertainment elements, such as gamification.
4. Educational: All intervention components that go beyond the simple feedback of generated PGHD (reflective), being attached to additional information that targets knowledge and skill enhancement, as well as knowledge testing. In contrast with process guiding, this category focuses on understanding and applying PGHD. They include the provision of newsletters, in-person counseling, remote coaching, educational podcasts, quizzes, and knowledge tests. Technical guidance on the generation and share of PGHD falls out of this category's scope (see point 2. Process guiding).

The integration and utilization of electronic PGHD varied across additional intervention components. In combination with reflective components, PGHD were mostly used for self-referencing, such as the visualization of progress over time, enabling users to track individual health goals. In the context of motivational components, PGHD were repeatedly utilized to enable social comparison, such as contrasting data to normative or peer-generated values, often generating certain social pressure for healthier lifestyles. When integrated with educational components, PGHD were used to inform and provide individualized recommendations and counseling, aligned to the progress and capabilities of individual participants or participant subgroups. Finally, in combination with process guiding components, PGHD were key to identifying individual challenges, allowing for tailored and problem-focused support, while ensuring that adherence to intervention guidelines (eg, dietary or exercise plans) was monitored.

The Role of Health Care Providers

Less than half of the literature described the role of health care providers in the implementation (73/183, 39.9%) of interventions and only 30.6% (56/183) involved providers that had clearly designated PGHD-related responsibilities. The remaining proportion of the literature (110/183, 60.1%) addressed predominantly automatic programs. The proportion of studies without provider involvement was larger across all prevention areas, except for those on cardiometabolic health and breastfeeding promotion. Health care providers included an array of professionals, including physicians, nurses, dieticians, psychologists, health consultants, fitness experts, and trainers. Our thematic analysis identified 2 main clusters of PGHD-related provider roles. The first role (5/55, 9%) is that of a supporter, including the prompting, overseeing, and motivating of PGHD use, which was primarily found in weight control, nutrition and well-being interventions. The second role (34/55, 62%) is that of a reviewer, consisting of analyzing PGHD to inform counseling, personalize advise, conduct remote monitoring, and complement medical data, which was common in weight control, nutrition, cardiometabolic health, and substance use prevention. In 31% (17/55) of the studies, mainly in weight control, cardiometabolic health, and substance use prevention, providers held both roles simultaneously. In addition, we identified that provider-consumer interactions predominantly occurred remotely (36/73, 49%), either via the PGHD tool itself (eg, data collection website) or through other supporting channels (eg, email), both in a synchronous (eg, telephone) or asynchronous fashion (eg, forums). In-person interactions were less common (14/73, 19%) and more often combined with remote elements (17/73, 23%). One study (1/73, 1%) involved no direct interaction with consumers, whereas 5 (5/73, 7%) lacked clear interaction descriptions. Our findings additionally suggested that the involvement of health care providers was linked to the previously described PGHD aims, as one of those, namely the aim of enhanced remote monitoring, inevitably relies on data review by a provider. Table 3 provides the distribution of provider roles across prevention areas.

Patient-Generated Health Data Implications: Health Impact and Equity

Assessed prevention-relevant outcomes were broadly categorized into: (1) vitals and body composition measures (eg, body mass index, blood pressure, blood glucose, and heart rate), (2) behavioral change (eg, physical activity, eating habits, and lifestyle factors), and (3) knowledge change (eg, health literacy and awareness). About a quarter of the identified literature (47/183, 25.7%) provided indications on the potential impact

of PGHD on preventive outcomes. These studies either had PGHD as a distinct or single component in one of their intervention arms (13/47, 28%), or as a part of multicomponent interventions (34/47, 72%), with sections that explored the associations between PGHD (eg, adherence to data collection) and outcomes. The majority explored implications on vitals and body composition-related outcomes (37/47, 79%). Most of those studies reported statistically significant beneficial trends (n=27), followed by nonsignificant effects (n=8) and mixed results (n=4). Outcomes in health behavior were less commonly addressed (15/47, 32%) and provided no clear tendencies, with an equal number of studies providing statistically significant beneficial (n=4) and nonsignificant (n=4) trends, as well as a relatively large proportion of unclear or mixed results (n=3). Health knowledge outcomes were the least commonly (2/47, 4%) addressed, with one study reporting nonsignificant associations between PGHD and health knowledge and one reporting mixed results. Most of these studies included active (27/47, 57%) and partially active consumers (8/47, 17%), whereas only one study entailed passive consumers (1/47, 2%). For the studies with active and partially active user engagement, a proportionally equal number of them reported statistically significant, mixed, and nonstatistically significant results. One study that included passive consumers did not provide enough information to be meaningfully compared.

A larger proportion of the literature (98/183, 53.6%) addressed interventions with multiple components and did not entail analyses on the relationship between PGHD components and prevention outcomes. Although their results could not be directly linked to PGHD, the overall picture suggested beneficial trends, with 23% (22/98) providing almost entirely positive results. Mixed results were indicated by 69% (68/98) of studies, almost all of which included at least one significantly positive outcome. A smaller proportion of interventions (8/96, 8%) did not identify beneficial effects at all. The remaining part of the literature (38/183, 20.8%) focused on feasibility and usability results instead, which is not reported in further detail here.

Considering equity as an important outcome for all health interventions, we extracted information linked to implications for subgroups that are commonly divided by health inequalities, as defined by the Cochrane Equity Group [23]. About 46.5% of studies (85/183) addressed equity by referring to the digital divide and literacy inequalities, by addressing the limitations of homogeneous study samples that primarily consisted of advantaged subgroups (eg, white, highly educated), by focusing on underserved populations, and by exploring patterns across sociodemographic variables. Approximately 7.7% of the literature (14/183) provided detailed analyses across subgroups divided by sex (12/14), age (6/14), race (4/14), education (4/14), income (2/14), and place of residence (1/14). Most of those indicated either no or unclear differential effects, whereas 2 indicated better intervention effects for younger participants, one for white non-Hispanic individuals and one for higher educated participants.

Barriers and Facilitators of Electronic Patient-Generated Health Data Use

About 89.6% (164/183) of studies provided information on potential barriers and facilitators of electronic PGHD. Both barriers and facilitators were clustered around 5 recurring themes: (1) data collection and use, (2) technical and design considerations, (3) ethics, safety, and trust, (4) user characteristics, and (5) context and content. Data collection and use (127/164, 77.4%) addressed the levels of ease, difficulty, and burden of electronic PGHD generation, the adaptability of data collection to user needs, and associated resource demands (eg, time, costs). Technicalities and design (84/164, 51.2%) covered the functional maturity of PGHD technology, the facilitating role of mobile and interoperable devices, as well as the importance of dynamic, user-appealing, and simple designs. Ethics, safety, and trust (55/164, 33.5%) entailed barriers and facilitators around privacy, trustworthiness, credibility, and reliability. The category of user characteristics (72/164, 43.9%) highlighted consumer-related elements, such as digital literacy, knowledge, sociodemographic determinants, and overall attitudes toward PGHD technologies. Finally, the last category of content and context (148/164, 90.2%) included elements around contextual resources, such as PGHD support and interaction with providers. It additionally addressed the role of technology and intervention content, such as the combination of PGHD with other behavior change communication techniques. Textboxes 3 and 4 provide a list of identified barriers and facilitators across those 5 themes.

Textbox 3. Barriers of electronic patient-generated health data and the number of studies reporting barriers in each domain.

Data collection and use (n=49):

- Burdensome data collection
- Inflexible data entry
- Retrospective data entry: incentive to manipulate data
- Unstructured data: information overload
- Automatized recording: feeling of no control over data
- Costly

Technicalities and design (n=39):

- Immature or nonfunctional
- Unappealing design
- Nonuser-friendly functions

Ethics, safety, and trust (n=32):

- Privacy and security concerns
- Nontrustworthy patient-generated health data (PGHD) tools and data
- Sociocultural resistances
- Low-quality and unreliable PGHD

User characteristics (n=38):

- Low digital literacy and no previous experience
- Negative attitudes toward PGHD
- Mismatch with daily life routines
- Nonperceived usefulness
- Sociodemographics (eg, young age, low education)

Content and context (n=44):

- Missing data interpretation and general support
- Missing in-person contact
- Missing or too frequent reminders
- Missing provider resources to evaluate PGHD
- Missing (financial) incentives or rewards
- Missing or insensitive feedback
- Unrealistic goals
- Nonengaging environment (eg, no social support)

Textbox 4. Facilitators of electronic patient-generated health data and the number of studies reporting facilitators in each domain.

Data collection and use (n=78):

- Simple and low-effort data collection
- Highly flexible data entry
- Retrospective data entry: incentive to correct data
- Time-efficient and intuitive data output
- Automatized recording: noninterference with daily life
- Free or low cost

Technicalities and design (n=45):

- Technically mature and interoperable
- User-engaging and appealing design
- Dynamic design: interactive and modifiable
- Mobile

Ethics, safety, and trust (n=23):

- Credible patient-generated health data (PGHD) tools
- Trustworthy, reliable, and complete PGHD
- Processes that do not invade privacy

User characteristics (n=34):

- Digital literacy and previous experience
- Preexisting motivation and readiness to use PGHD
- Self-efficacy
- Perceived PGHD usefulness and relevance

Content and context (n=104):

- Available guidance and support
- Available human interaction
- Sensitive reminders that do not disturb
- Data visualizations and summaries
- Motivating rewards and (financial) incentives
- Immediate, sensitive, and motivating feedback
- Realistic goal setting
- Social support (eg, peer interactions)
- Content and context personalization
- Enabled data access, ownership, and control
- Fun elements (eg, gamification)
- Novel elements (eg, geofence triggered support)

Discussion

Overview

Our review described a large and dynamically emerging volume of the literature on the use of electronic PGHD for primary disease prevention and health promotion purposes. Beyond quantity, the literature manifested large methodological and thematic heterogeneity, adding to the topic's conceptual complexity. Our results enabled the development of an enriched conceptual framework (Figure 2) and thus a better understanding of the process between generating PGHD and finally utilizing them for preventive and health promoting action.

Principal Findings and Comparison With Previous Work

The identified literature predominantly focused on weight control, through physical activity and nutrition, which was consistent with previous reviews that addressed digital health interventions across prevention areas [36-38]. In line with the variety of existing digital approaches to primary disease prevention [39-40], our results indicated that electronic PGHD target multiple dimensions, including: (1) health knowledge and self-awareness, (2) behavior change, (3) healthy environments, and (4) remote monitoring. Overall, we identified 4 types of prevention-targeted interventions entailing electronic PGHD: (1) automatic, single component, (2) automatic multicomponent interventions, (3) single component, and (4) multicomponent interventions that are not fully automatic, including health care provider involvement (Figure 2). A single component denotes that PGHD is the main and only prevention element, whereas a multicomponent describes interventions with at least 1 additional non-PGHD element.

Our thematic analysis identified certain recurring patterns of PGHD generation. We broadly classified consumer roles as passive, partially active, and fully active and identified that the proportions of these vary across prevention areas. Acknowledging that consumer roles are closely linked to PGHD types, we found that certain prevention areas are being dominated by 1 or 2 types of PGHD. On one hand, weight control, alcohol and smoking prevention, and overall health and well-being seemed to be mostly addressed by technologies that require the manual collection of textual or numerical data, while on the other hand, cardiometabolic disease prevention was primarily addressed by a combination of PGHD types that require a mix of active and passive data generation. In contrast, entirely passive data generation was only identified for weight control, overall health and well-being, cardiometabolic health, and healthy aging. Although not focused on prevention, a review by Vagesna et al [41] identified similar patterns, where weight was monitored by computerized systems (manual data entry) and metabolic conditions by a combination of multiple technologies.

Considering that the sophistication and reliability of PGHD technology varies across prevention areas, these patterns are expected. To be adequately targeted and well-informed, prevention often requires very specific consumer action and PGHD input. On one hand, for certain areas, such as addiction prevention or dietary intake, this input is entirely behavioral and not easily captured automatically. This includes the exact number of smoked cigarettes and consumed alcohol drinks, the type of consumed drinks, the percentage of alcohol content in each drink or the portions of consumed meals, all of which currently cannot be reliably or cost effectively collected by sensor-based devices, while on the other hand, sophisticated and highly commercialized fitness trackers are increasingly being improved to reliably capture certain activities and bodily functions, such as physical exercise and heart rate. Exercise-based weight loss, well-being promotion (eg, sleep quality), and healthy aging (eg, fall prevention) are prevention areas in which such devices can be applied to, which explains the prevalence of passive PGHD generation. In between the two extremes, there are prevention approaches that inherently require combinations of measures, such as in diabetes prevention (eg, dietary intake and physical activity), which in turn allow for a partially active and partially passive generation of PGHD.

Linking consumer roles to identified barriers and facilitators suggests some conflicting dynamics. Passive PGHD generation might be less burdensome, but may also lead to lower consumer engagement. Conversely, active generation involves more effort but may simultaneously trigger higher user motivation. In their review on wearable monitoring technology, Baig et al [42] described part of these challenges, such as that passive application of technology might counteract user engagement and acceptance. Another review linked the intensity of consumer responsibilities to intervention effectiveness, reporting that programs with active consumers (eg, manual data input) were more successful [43]. To fully understand the most effective use of electronic PGHD for prevention, these patterns suggest that is important to further explore the interactions between PGHD types and their demands on consumers.

The relatively large proportion of studies that described automatic prevention systems, constituting 60.1% (110/183) of the identified literature, underlines a broader trend toward low-recourse and efficiency-driven care [44]. As expected, this was not the case for areas that traditionally rely on close patient-provider relationships, such as cardiometabolic health and breastfeeding. The remaining 39.9% (73/183) of the literature indicated 2 main health care provider clusters: (1) supporting PGHD collection and subsequently, (2) reviewing or using data for preventive practice. If fully automatized, those tasks were transferred to consumers, or the technology itself. A scoping review on the preventive use of smart devices by Petit and Cambon [45] described consumer responsibilities as a key literature aspect, in which patients gained expert roles and became equal agents of their own health care. Interalia, this has potential implications on data interpretation that we have identified as a recurring concept across studies. Previous research demonstrated the importance of PGHD interpretation and its wider implications [46], with missing sense-making skills and the fear of self-interpretation mentioned as key challenges to personal health information use. Our findings reflected that importance, as a major proportion of the literature directly or indirectly addressed interpretability by describing various approaches to fostering correct data understanding.

Most identified studies integrated electronic PGHD within multicomponent interventions, either complementing or facilitating other intervention components (eg, enabling self-reflection, facilitating social comparison, informing counseling, and directing guidance). A systematic review on the use of technology for weight reduction identified a similar trend, with 19 out of 27 studies combining PGHD with counseling feedback [47]. These trends highlight the value of PGHD for complex digital prevention, especially when not entailing health care providers. In addition, a previous scoping review by Lentferink et al identified that using PGHD to personalize intervention components (eg, goal setting) seemed to be a distinct element of highly effective studies [43].

Summarizing the results of single-component interventions, exploratory analyses (eg, associations between PGHD and health outcomes), and overall effects of multicomponent interventions, the overall directions suggest a predominantly positive PGHD impact on prevention. These trends are expected, considering the existing evidence on the association between monitoring one's own health and preventive outcomes [48-50]. On the contrary, trends on the impact of PGHD on health inequity gaps were not as clear, as only 7.7% (14/183) of the literature reported on differential effects across subgroups. The limited information and methodological ambiguity around subgroup analyses [51] did not allow for confident equity statements. Equity reviews from other fields, such as population and primary-care-based physical activity interventions reported on similar challenges, with only 19% (17/87) and 14.0% (24/171) of primary studies providing information on differential distribution of effects [52,53]. This is problematic, considering the potential implications of the digital divide, which is transforming from a divide on technology access to a divide on digital literacy and skills [54]. If we fail to adequately address equity, we risk limiting the preventive benefits of technology to those few who have the resources and skills to use them appropriately. Nonetheless, it is encouraging that almost half of the literature directly or indirectly mentioned equity, which indicates that many researchers are aware of the importance of ethics in digital health research.

Finally, the existence of conflicting barriers and facilitators highlighted the currently emerging nature and potential knowledge gaps of the topic. When do reminders become disturbances in one's daily life and when are they the key to ensuring prevention adherence? Are automatic and simple data collection methods preferred by consumers because of being less burdensome, or do they counteract user engagement and motivation? Do financial rewards act purely as incentives to collect data and adhere to preventive guidelines, or could they become incentives for data falsification? Are PGHD tools that allow for retrospective data entries beneficial because of added flexibility, or do they add to the risk of data manipulation? Although these uncertainties may be indicators of an emerging topic that requires more research, they might also be the result of electronic health and prevention complexity. Neither digitalization nor prevention are static or fixed phenomena [55]. They ultimately depend on interactions and contexts, for which single answers may be difficult to find. Nonetheless, the richness and range of identified barriers and facilitators indicate that the preventive use of electronic PGHD is sensitive to many factors, be it the way data are collected, the context in which they are collected, the personal characteristics of users, as well as ethical and technological considerations.

Limitations

Despite our rigorous methodological approach, our study is subjected to some limitations. First, although PGHD may be key throughout the continuum of care, our review's scope is restricted to primary disease prevention and health promotion. As such, our overall findings might not be applicable to the domains of treatment, disease management, and rehabilitation. This scope is narrower than defined in our protocol and has been chosen for practical and conceptual reasons. Retaining a broader scope would have led to an unmanageable volume of the literature and challenges in meaningfully synthesizing the results. Second, the chosen definition of electronic PGHD, which emphasized the aspect of patient's control, led to the exclusion of standardized and more provider-driven approaches. Broadening the definition might provide a better understanding of PGHD-based prevention within health care contexts and the interaction of such data streams with health care provider infrastructures. Third, the variety and evolvement of definitions and terms to describe PGHD, as well as prevention, might have led to missing out a few terms and the associated literature. To compensate for that, we conducted thorough hand searches, reference list screenings, and author consultations. Finally, our scoping methodology and heterogeneous output did not allow for robust synthesis and comparison of effects.

Implications and Future Research

Overall Implications

The patterns we identified may support users, patients, and providers in understanding the complexity of utilizing electronic PGHD for prevention purposes. Beyond technical maturity, providers need to consider the wider implications that data collection might have on patients and consumers, such as its interference with daily living, personal beliefs, and digital literacy. Users and providers need to be sensitive to ethical and trust concerns, while ensuring that the PGHD environments are motivating and supportive enough to facilitate adherence and successful prevention.

Future Research

Scoping reviews are often conducted to assess the feasibility of conducting a full systematic review [19]. On the basis of the large and very heterogeneous literature volume, we believe that conducting a full systematic review, while retaining a similarly broad scope, will be conceptually difficult. We, therefore, suggest (1) a narrower scope, (2) a careful *a priori* consideration of a PGHD definition, as that differed across the literature and may impact the review's results, (3) a careful selection of search terms, and (4) a preparatory literature scan to identify all terms appropriate to the chosen scope, as the terminology is vast and evolving. Beyond systematic reviews, future research should target evidence on best possible combinations of electronic PGHD with other behavioral change techniques (eg, feedback, goal setting, and peer interaction). Finally, future research should aim to capture how barriers and facilitators vary across contexts, while addressing the wider implications of PGHD-based prevention on the functioning of health systems and health equity.

Conclusions

Our review provides a comprehensive picture of the literature on electronic PGHD use for primary disease prevention and health promotion purposes, enabling a broader identification of processes and patterns. The high heterogeneity in the scope and content of identified studies underlines the topic's emerging nature. This is reflected by the variety of identified PGHD-generating technologies, resulting in diverse data types and different consumer responsibilities. Utilizing electronic PGHD to prevent disease and promote health is a complex

matter. In the literature, this complexity arises from electronic PGHD being mostly integrated into multicomponent and automatized interventions, limiting our ability to assess their individual preventive impact, and underlining an overall trend toward larger consumer responsibility. The broad set of identified barriers and facilitators, some being conflicting, highlights the need for a holistic understanding of such enabling factors, as well as for a stronger focus on ethical challenges, which is currently lacking.

Acknowledgments

The authors would like to thank Shannon Hubbs for acting as a third reviewer. The salary of VN was kindly provided by the Béatrice Ederer-Weber Foundation. The sponsor was not involved in any phase of this study. No further funding was received. The authors would like to thank Dr med Andreas Burkhard, Maria Giovanna Caruso, and Quentin Leonard Pannier for critically reviewing and providing feedback on the final manuscript. All authors provided submission approval for the final manuscript version.

Authors' Contributions

VN contributed to the conceptualization and implementation of the study. VN wrote and edited the manuscript. MM and MAP contributed to the conceptualization of the study, supervised the entire process, and critically edited the manuscript. FE contributed to the design of the study, provided regular feedback, and critically edited the manuscript. PL contributed to the implementation of study and critically edited the manuscript. The main screening and review procedures were conducted by VN and PL.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Protocol deviations and justifications.
[PDF File (Adobe PDF File), 18 KB - jmir_v21i10e13320_app1.pdf]

Multimedia Appendix 2
Search strategy and utilized keywords.
[PDF File (Adobe PDF File), 92 KB - jmir_v21i10e13320_app2.pdf]

Multimedia Appendix 3
Adapted conceptual framework of electronic patient-generated health data flow.
[PDF File (Adobe PDF File), 208 KB - jmir_v21i10e13320_app3.pdf]

Multimedia Appendix 4
List of studies excluded at full-text appraisal, with reasons.
[PDF File (Adobe PDF File), 270 KB - jmir_v21i10e13320_app4.pdf]

Multimedia Appendix 5
List and characteristics of included studies.
[PDF File (Adobe PDF File), 346 KB - jmir_v21i10e13320_app5.pdf]

Multimedia Appendix 6
Extracted study characteristics.
[XLSX File (Microsoft Excel File), 38 KB - jmir_v21i10e13320_app6.xlsx]

Multimedia Appendix 7
Participant characteristics.
[PDF File (Adobe PDF File), 113 KB - jmir_v21i10e13320_app7.pdf]

Multimedia Appendix 8
List of included studies, grouped by prevention area.
[XLSX File (Microsoft Excel File), 27 KB - jmir_v21i10e13320_app8.xlsx]

Multimedia Appendix 9
Patient-generated health data tools and their functionalities.
[PDF File (Adobe PDF File), 13 KB - jmir_v21i10e13320_app9.pdf]

References

1. Castle-Clarke S, Imison C. The Nuffield Trust. 2016. The Digital Patient: Transforming Primary Care? Research Report URL: https://www.nuffieldtrust.org.uk/files/2017-06/1497259872_nt-the-digital-patient-web-corrected-p46-.pdf [accessed 2019-01-03] [WebCite Cache ID 759ZBozlK]
2. Car J, Tan WS, Huang Z, Sloot P, Franklin BD. eHealth in the future of medications management: personalisation, monitoring and adherence. BMC Med 2017 Apr 5;15(1):73 [FREE Full text] [doi: 10.1186/s12916-017-0838-0] [Medline: 28376771]
3. Federation of American Scientists. 2008. Disruptive Civil Technologies: Six Technologies with Potential Impacts on US Interests out to 2025 URL: https://fas.org/irp/nic/disruptive.pdf [accessed 2019-01-04] [WebCite Cache ID 75AcJC9PV]
4. Maltseva K, Lutz C. A quantum of self: a study of self-quantification and self-disclosure. Comput Hum Behav 2018 Apr;81:102-114 [FREE Full text] [doi: 10.1016/j.chb.2017.12.006]
5. Hsueh PS, Dey S, Das S, Wetter T. Making sense of patient-generated health data for interpretable patient-centered care: the transition from 'more' to 'better'. Stud Health Technol Inform 2017;245:113-117. [Medline: 29295063]
6. Shapiro M, Johnston D, Wald J, Mon D. HealthIT.gov. 2012. Patient-Generated Health Data: White Paper URL: https://www.healthit.gov/sites/default/files/rti_pghd_whitepaper_april_2012.pdf [accessed 2019-01-04] [WebCite Cache ID 75AdcD3ei]
7. Fox S, Sydney J. Pew Research Center. 2009. The Social Life of Health Information URL: http://www.pewinternet.org/2009/06/11/the-social-life-of-health-information/ [accessed 2019-01-04] [WebCite Cache ID 75AdoMjI3]
8. Hartzler AL, McDonald DW, Park A, Huh J, Weaver C, Pratt W. Evaluating health interest profiles extracted from patient-generated data. AMIA Annu Symp Proc 2014;2014:626-635 [FREE Full text] [Medline: 25954368]
9. Visuri A, van Berkel N, Luo C, Goncalves J, Ferreira D, Kostakos V. Challenges of Quantified-Self: Encouraging Self-Reported Data Logging During Recurrent Smartphone Usage. In: Proceedings of the 31st British Computer Society Human Computer Interaction Conference. 2017 Presented at: HCI'17; July 3-6, 2017; Sunderland, UK. [doi: 10.14236/ewic/hci2017.81]
10. Maddison R, Gemming L, Monedero J, Bolger L, Belton S, Issartel J, et al. Quantifying human movement using the Movn smartphone app: validation and field study. JMIR Mhealth Uhealth 2017 Aug 17;5(8):e122 [FREE Full text] [doi: 10.2196/mhealth.7167] [Medline: 28818819]
11. Vangeepuram N, Mayer V, Fei K, Hanlen-Rosado E, Andrade C, Wright S, et al. Smartphone ownership and perspectives on health apps among a vulnerable population in East Harlem, New York. Mhealth 2018;4:31 [FREE Full text] [doi: 10.21037/mhealth.2018.07.02] [Medline: 30221166]
12. Guariguata L, Whiting DR, Hambleton I, Beagley J, Linnenkamp U, Shaw JE. Global estimates of diabetes prevalence for 2013 and projections for 2035. Diabetes Res Clin Pract 2014 Mar;103(2):137-149. [doi: 10.1016/j.diabres.2013.11.002] [Medline: 24630390]
13. Sharon T. Self-tracking for health and the quantified self: re-articulating autonomy, solidarity, and authenticity in an age of personalized healthcare. Philos Technol 2016 Apr 18;30(1):93-121. [doi: 10.1007/s13347-016-0215-5]
14. Lv N, Xiao L, Simmons ML, Rosas LG, Chan A, Entwistle M. Personalized hypertension management using patient-generated health data integrated with electronic health records (EMPOWER-H): six-month pre-post study. J Med Internet Res 2017 Sep 19;19(9):e311 [FREE Full text] [doi: 10.2196/jmir.7831] [Medline: 28928111]
15. Wijsman CA, Westendorp RG, Verhagen EA, Catt M, Slagboom PE, de Craen SJ, et al. Effects of a web-based intervention on physical activity and metabolism in older adults: randomized controlled trial. J Med Internet Res 2013 Nov 6;15(11):e233 [FREE Full text] [doi: 10.2196/jmir.2843] [Medline: 24195965]
16. HealthIT.gov. 2014. Patient-Generated Health Data: Fact Sheet URL: https://www.healthit.gov/sites/default/files/patient_generated_data_factsheet.pdf [accessed 2019-01-04] [WebCite Cache ID 75AjrqBwz]
17. Sands DZ, Wald JS. Transforming health care delivery through consumer engagement, health data transparency, and patient-generated health information. Yearb Med Inform 2014 Aug 15;9:170-176 [FREE Full text] [doi: 10.15265/IY-2014-0017] [Medline: 25123739]
18. Hull S. Patient-generated health data foundation for personalized collaborative care. Comput Inform Nurs 2015 May;33(5):177-180. [doi: 10.1097/CIN.0000000000000159] [Medline: 25974360]
19. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodol 2005 Feb 23;8(1):19-32. [doi: 10.1080/1364557032000119616]
20. Levac D, Colquhoun H, O'Brien KK. Scoping studies: advancing the methodology. Implement Sci 2010 Sep 20;5:69 [FREE Full text] [doi: 10.1186/1748-5908-5-69] [Medline: 20854677]
21. Nittas V, Mütsch M, Ehrler F, Puhan MA. Electronic patient-generated health data to facilitate prevention and health promotion: a scoping review protocol. BMJ Open 2018 Aug 10;8(8):e021245 [FREE Full text] [doi: 10.1136/bmjopen-2017-021245] [Medline: 30099392]

22. Cohen DJ, Keller SR, Hayes GR, Dorr DA, Ash JS, Sittig DF. Integrating patient-generated health data into clinical care settings or clinical decision-making: lessons learned from project healthdesign. JMIR Hum Factors 2016 Oct 19;3(2):e26 [FREE Full text] [doi: 10.2196/humanfactors.5919] [Medline: 27760726]
23. Welch V, Petticrew M, Petkovic J, Moher D, Waters E, White H, PRISMA-Equity Bellagio Group. Extending the PRISMA statement to equity-focused systematic reviews (PRISMA-E 2012): explanation and elaboration. Int J Equity Health 2015 Oct 8;14:92 [FREE Full text] [doi: 10.1186/s12939-015-0219-2] [Medline: 26450828]
24. Daudt HM, van Mossel C, Scott SJ. Enhancing the scoping study methodology: a large, inter-professional team's experience with Arksey and O'Malley's framework. BMC Med Res Methodol 2013 Mar 23;13:48 [FREE Full text] [doi: 10.1186/1471-2288-13-48] [Medline: 23522333]
25. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Ann Intern Med 2009 Aug 18;151(4):264-9, W64. [doi: 10.7326/0003-4819-151-4-200908180-00135] [Medline: 19622511]
26. Kerr DA, Harray AJ, Pollard CM, Dhaliwal SS, Delp EJ, Howat PA, et al. The connecting health and technology study: a 6-month randomized controlled trial to improve nutrition behaviours using a mobile food record and text messaging support in young adults. Int J Behav Nutr Phys Act 2016 Apr 21;13:52 [FREE Full text] [doi: 10.1186/s12966-016-0376-8] [Medline: 27098449]
27. Beasley JM, Riley WT, Davis A, Singh J. Evaluation of a PDA-based dietary assessment and intervention program: a randomized controlled trial. J Am Coll Nutr 2008 Apr;27(2):280-286. [doi: 10.1093/aje/163.suppl_11.S38] [Medline: 18689560]
28. Gotsis M, Wang H, Spruijt-Metz D, Jordan-Marsh M, Valente TW. Wellness partners: design and evaluation of a web-based physical activity diary with social gaming features for adults. JMIR Res Protoc 2013 Mar 1;2(1):e10 [FREE Full text] [doi: 10.2196/resprot.2132] [Medline: 23611986]
29. Youm S, Liu S. Development healthcare PC and multimedia software for improvement of health status and exercise habits. Multimed Tools Appl 2015 Oct 19;76(17):17751-17763. [doi: 10.1007/s11042-015-2998-2]
30. Cordeiro F, Bales E, Cherry E, Fogarty J. Rethinking the Mobile Food Journal: Exploring Opportunities for Lightweight Photo-Based Capture. In: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems. 2015 Presented at: CHI'15; April 18-23, 2015; Seoul, Republic of Korea p. 3207-3216. [doi: 10.1145/2702123.2702154]
31. Gaggioli A, Cipresso P, Serino S, Campanaro DM, Pallavicini F, Wiederhold BK, et al. Positive technology: a free mobile platform for the self-management of psychological stress. Annu Rev Cyber Ther Telemed 2014;12:25-29 [FREE Full text]
32. Bond DS, Thomas JG, Raynor HA, Moon J, Sieling J, Trautvetter J, et al. B-MOBILE--a smartphone-based intervention to reduce sedentary time in overweight/obese individuals: a within-subjects experimental trial. PLoS One 2014;9(6):e100821 [FREE Full text] [doi: 10.1371/journal.pone.0100821] [Medline: 24964010]
33. King AC, Hekler EB, Grieco LA, Winter SJ, Sheats JL, Buman MP, et al. Harnessing different motivational frames via mobile phones to promote daily physical activity and reduce sedentary behavior in aging adults. PLoS One 2013;8(4):e62613 [FREE Full text] [doi: 10.1371/journal.pone.0062613] [Medline: 23638127]
34. Slootmaker SM, Chinapaw MJ, Schuit AJ, Seidell JC, van Mechelen W. Feasibility and effectiveness of online physical activity advice based on a personal activity monitor: randomized controlled trial. J Med Internet Res 2009 Jul 29;11(3):e27 [FREE Full text] [doi: 10.2196/jmir.1139] [Medline: 19674956]
35. Bove AA, Santamore WP, Homko C, Kashem A, Cross R, McConnell TR, et al. Reducing cardiovascular disease risk in medically underserved urban and rural communities. Am Heart J 2011 Mar;161(2):351-359. [doi: 10.1016/j.ahj.2010.11.008] [Medline: 21315219]
36. Prochaska JJ, Prochaska JO. A review of multiple health behavior change interventions for primary prevention. Am J Lifestyle Med 2011 May;5(3):208-221 [FREE Full text] [doi: 10.1177/1559827610391883] [Medline: 24358034]
37. Kampmeijer R, Pavlova M, Tambor M, Golinowska S, Groot W. The use of e-health and m-health tools in health promotion and primary prevention among older adults: a systematic literature review. BMC Health Serv Res 2016 Sep 5;16(Suppl 5):290 [FREE Full text] [doi: 10.1186/s12913-016-1522-3] [Medline: 27608677]
38. Kohl LF, Crutzen R, de Vries NK. Online prevention aimed at lifestyle behaviors: a systematic review of reviews. J Med Internet Res 2013 Jul 16;15(7):e146 [FREE Full text] [doi: 10.2196/jmir.2665] [Medline: 23859884]
39. Coughlin SS, Prochaska JJ, Williams LB, Besenyi GM, Heboyan V, Goggans DS, et al. Patient web portals, disease management, and primary prevention. Risk Manag Healthc Policy 2017;10:33-40 [FREE Full text] [doi: 10.2147/RMHP.S130431] [Medline: 28435342]
40. Ahmad FS, Tsang T. Diabetes prevention, health information technology, and meaningful use: challenges and opportunities. Am J Prev Med 2013 Apr;44(4 Suppl 4):S357-S363. [doi: 10.1016/j.amepre.2012.12.020] [Medline: 23498299]
41. Vegesna A, Tran M, Angelaccio M, Arcona S. Remote patient monitoring via non-invasive digital technologies: a systematic review. Telemed J E Health 2017 Jan;23(1):3-17 [FREE Full text] [doi: 10.1089/tmj.2016.0051] [Medline: 27116181]
42. Baig MM, GholamHosseini H, Moqeem AA, Mirza F, Lindén M. A systematic review of wearable patient monitoring systems - current challenges and opportunities for clinical adoption. J Med Syst 2017 Jul;41(7):115. [doi: 10.1007/s10916-017-0760-1] [Medline: 28631139]

43. Lentferink AJ, Oldenhuis HK, de Groot M, Polstra L, Velthuijsen H, van Gemert-Pijnen JE. Key components in ehealth interventions combining self-tracking and persuasive ecoaching to promote a healthier lifestyle: a scoping review. J Med Internet Res 2017 Aug 1;19(8):e277 [FREE Full text] [doi: 10.2196/jmir.7288] [Medline: 28765103]
44. Kvedar J, Coye MJ, Everett W. Connected health: a review of technologies and strategies to improve patient care with telemedicine and telehealth. Health Aff (Millwood) 2014 Mar;33(2):194-199. [doi: 10.1377/hlthaff.2013.0992] [Medline: 24493760]
45. Petit A, Cambon L. Exploratory study of the implications of research on the use of smart connected devices for prevention: a scoping review. BMC Public Health 2016 Jul 11;16:552 [FREE Full text] [doi: 10.1186/s12889-016-3225-4] [Medline: 27401769]
46. Hartmann-Boyce J, Boylan AM, Jebb SA, Aveyard P. Experiences of self-monitoring in self-directed weight loss and weight loss maintenance: systematic review of qualitative studies. Qual Health Res 2019 Jan;29(1):124-134. [doi: 10.1177/1049732318784815] [Medline: 29984630]
47. Raaijmakers LC, Pouwels S, Berghuis KA, Nienhuijs SW. Technology-based interventions in the treatment of overweight and obesity: a systematic review. Appetite 2015 Dec;95:138-151. [doi: 10.1016/j.appet.2015.07.008] [Medline: 26165415]
48. Teasdale N, Elhussein A, Butcher F, Piernas C, Cowburn G, Hartmann-Boyce J, et al. Systematic review and meta-analysis of remotely delivered interventions using self-monitoring or tailored feedback to change dietary behavior. Am J Clin Nutr 2018 Feb 1;107(2):247-256 [FREE Full text] [doi: 10.1093/ajcn/nqx048] [Medline: 29529158]
49. Burke LE, Wang J, Sevick MA. Self-monitoring in weight loss: a systematic review of the literature. J Am Diet Assoc 2011 Jan;111(1):92-102 [FREE Full text] [doi: 10.1016/j.jada.2010.10.008] [Medline: 21185970]
50. Zheng Y, Klem ML, Sereika SM, Danford CA, Ewing LJ, Burke LE. Self-weighing in weight management: a systematic literature review. Obesity (Silver Spring) 2015 Mar;23(2):256-265 [FREE Full text] [doi: 10.1002/oby.20946] [Medline: 25521523]
51. Petticrew M, Tugwell P, Kristjansson E, Oliver S, Ueffing E, Welch V. Damned if you do, damned if you don't: subgroup analysis and equity. J Epidemiol Community Health 2012 Jan;66(1):95-98. [doi: 10.1136/jech.2010.121095] [Medline: 21652518]
52. Humphreys DK, Ogilvie D. Synthesising evidence for equity impacts of population-based physical activity interventions: a pilot study. Int J Behav Nutr Phys Act 2013 Jun 15;10:76 [FREE Full text] [doi: 10.1186/1479-5868-10-76] [Medline: 23768212]
53. Attwood S, van Sluijs E, Sutton S. Exploring equity in primary-care-based physical activity interventions using PROGRESS-Plus: a systematic review and evidence synthesis. Int J Behav Nutr Phys Act 2016 May 20;13:60 [FREE Full text] [doi: 10.1186/s12966-016-0384-8] [Medline: 27206793]
54. Neter E, Brainin E. eHealth literacy: extending the digital divide to the realm of health information. J Med Internet Res 2012 Jan 27;14(1):e19 [FREE Full text] [doi: 10.2196/jmir.1619] [Medline: 22357448]
55. Moller AC, Merchant G, Conroy DE, West R, Hekler E, Kugler KC, et al. Applying and advancing behavior change theories and techniques in the context of a digital health revolution: proposals for more effectively realizing untapped potential. J Behav Med 2017 Mar;40(1):85-98 [FREE Full text] [doi: 10.1007/s10865-016-9818-7] [Medline: 28058516]

Abbreviations

PGHD: patient-generated health data
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses

Edited by G Eysenbach; submitted 07.01.19; peer-reviewed by S Pekkarinen, S Mikles; comments to author 27.04.19; revised version received 04.06.19; accepted 19.08.19; published 14.10.19.

Please cite as:
Nittas V, Lun P, Ehrler F, Puhan MA, Mütsch M
Electronic Patient-Generated Health Data to Facilitate Disease Prevention and Health Promotion: Scoping Review
J Med Internet Res 2019;21(10):e13320
URL: http://www.jmir.org/2019/10/e13320/
doi:10.2196/13320
PMID:31613225

Original Paper

A Web-Based Intervention for Social Media Addiction Disorder Management in Higher Education: Quantitative Survey Study

Huseyin Dogan[1], PhD; Helmi Norman[2], PhD; Amen Alrobai[3], PhD; Nan Jiang[1], PhD; Norazah Nordin[2], PhD; Anita Adnan[4], PhD

[1]Bournemouth University, Bournemouth, United Kingdom

[2]Faculty of Education, Universiti Kebangsaan Malaysia, Bangi, Malaysia

[3]King Abdulaziz University, Jeddah, Saudi Arabia

[4]Doctorate Support Group, Kuala Lumpur, Malaysia

Corresponding Author:
Helmi Norman, PhD
Faculty of Education
Universiti Kebangsaan Malaysia
Jalan Reko
Bangi, 43600
Malaysia
Phone: 60 389216062
Email: helmi.norman@ukm.edu.my

Abstract

Background: Social media addiction disorder has recently become a major concern and has been reported to have negative impacts on postgraduate studies, particularly addiction to Facebook. Although previous studies have investigated the effects of Facebook addiction disorder in learning settings, there still has been a lack of studies investigating the relationship between online intervention features for Facebook addiction focusing on postgraduate studies.

Objective: In an attempt to understand this relationship, this study aimed to carry out an investigation on online intervention features for effective management of Facebook addiction in higher education.

Methods: This study was conducted quantitatively using surveys and partial least square-structural equational modeling. The study involved 200 postgraduates in a Facebook support group for postgraduates. The Bergen Facebook Addiction test was used to assess postgraduates' Facebook addiction level, whereas online intervention features were used to assess postgraduates' perceptions of online intervention features for Facebook addiction, which are as follows: (1) self-monitoring features, (2) manual control features, (3) notification features, (4) automatic control features, and (5) reward features.

Results: The study discovered six Facebook addiction factors (relapse, conflict, salience, tolerance, withdrawal, and mood modification) and five intervention features (notification, auto-control, reward, manual control, and self-monitoring) that could be used in the management of Facebook addiction in postgraduate education. The study also revealed that relapse is the most important factor and mood modification is the least important factor. Furthermore, findings indicated that notification was the most important intervention feature, whereas self-monitoring was the least important feature.

Conclusions: The study's findings (addiction factors and intervention features) could assist future developers and educators in the development of online intervention tools for Facebook addiction management in postgraduate education.

(J Med Internet Res 2019;21(10):e14834) doi:10.2196/14834

KEYWORDS

Facebook addiction; intervention features; postgraduate education; social media addiction; obsessive-compulsive disorder (OCD); PLS-SEM analysis

Introduction

Background

Addiction is usually associated with addictive behavior and compulsive engagements of stimuli, such as a psychoactive chemical (eg, alcohol and drugs), despite harmful consequences. Nevertheless, behavioral addiction related to nonconsumption of substances, such as digital addiction, has recently become a topic of much interest. To date, the only psychiatric disorders that have been formally recognized are pathological gambling and internet gaming disorder [1,2]. As such, there is an urgent need for further research in terms of behavioral addiction [3]. As social media has become an essential platform for online communication, several studies have investigated its behavioral effects on excessive usage. Although some researchers have addressed general digital and internet addiction [4] and its psychological effects on loneliness, anxiety, and depression [5], other researchers have focused on addiction of social networking sites (SNSs) such as Facebook [2,5].

As of June 2018, current statistics have indicated that Facebook is the most popular social network worldwide, with over 2 billion monthly active users. Despite its benefits for the Web-based social communication and content consumption, some users develop an excessive usage of Facebook, causing potential negative effects, termed as Facebook addiction disorder [2,5]. Facebook addiction is defined as an addictive behavior caused by an uncontrollable level of accessing and using Facebook, which negatively affects other face-to-face social activities, studies, jobs, interpersonal relationships, and physical health [6,7].

Facebook addiction disorder is categorized by psychological factors such as salience, tolerance, mood modification, relapse, withdrawal, and conflict [8]. Salience is related to the mental state of continuously thinking about Facebook, whereas tolerance is related to the tolerance level of Facebook usage (eg, increase the time spent on Facebook to reach to the same effect that was initially experienced using Facebook). Mood modification is associated with whether Facebook affects current moods of the user, and relapse is linked with failed attempts of Facebook usage reduction. Meanwhile, withdrawal and conflict are related to negative conditions and effects because of failure in accessing Facebook, in which withdrawal is associated with negative conditions such as becoming restless because of failure in accessing Facebook, whereas conflict is linked with negative effects such as Facebook causing negative impacts on individuals' current academic or professional life [2,9].

Previous research has revealed that Facebook addiction have caused negative psychological effects such as emotional problems, relational problems, health-related problems, and performance problems [9]. In terms of emotional problems, Facebook addiction has been revealed to cause negative mood alterations such as depression and anxiety [10], development of deficient self-regulation [11], as well as task avoidance and procrastination [12]. With regard to relational problems, Facebook addicts have experienced negative relationships in terms of family conflicts, impaired concentration at work or school, and problematic peer relationships, thus contributing to interpersonal relationship detriment [9,13]. With regard to health-related problems, Facebook addiction has also been associated with sleep difficulties such as insomnia and somatic problems as well as poorer sleep quality [9,14]. Meanwhile, for performance problems, addiction to Facebook has caused job losses and negative effects of self-reported work performance [9,14].

Facebook addiction has also been reported to affect higher education studies. In Turkey, a study on the effect of Facebook addiction on gender was carried out with 1257 Turkish university students [15]. The study's findings revealed that male students had higher SNS addiction levels as compared with female students. In Poland, Facebook addiction was studied among 1157 students [6]. They discovered that Facebook addiction among Polish students was related to higher extraversion, narcissism, loneliness, and social anxiety and lower general self-efficacy. They also discovered that Facebook addiction was further related to impoverished well-being that included impaired general health, decreased sleep quality, and higher perceived stress. In the United States, an investigation was conducted with 274 university students in a statistics course, in which they examined the time distortion of social media addiction in at-risk students by intervention strategies such as prevention of Facebook use and self-control strategies [16]. They discovered that the at-risk group showed a significant upward time estimate bias when positively correlated with Facebook addiction scores. In Malaysia, Facebook addiction motives were studied with 380 postgraduates and undergraduates and the study revealed that motives that contribute to addiction are factors such as social interaction, passing time, entertainment, companionship, and communication. In another study, 441 Malaysian students were assessed on their addiction to Facebook [11]. They found out that factors such as religion, level of income, ego strength, and locus of control do not show significant influence on the risk of Facebook addiction, whereas time spent on Facebook contributed to higher addiction levels [17]. Nevertheless, these studies only mostly focused on motives of Facebook addiction level, and this shows that there is a lack of studies on the Web-based intervention systems for Facebook addiction.

Objectives

Therefore, the main aim of this study was to investigate features of the Web-based intervention systems on management of Facebook addiction in postgraduate education. Considering the lack of knowledge on the development of the Web-based intervention features for Facebook addiction, the study was exploratory in nature [2,18]. The second aim was to investigate which intervention feature was most and least important for the management of addiction to Facebook. This could contribute to a better understanding of addiction prevention and therapeutic interventions of Facebook addiction. In the study, Web-based intervention features focused on features such as manual monitoring, manual limit, automatic notification, automatic limit, and automatic reward for learners to manage Facebook addiction. These features can be linked to an upcoming tool that is going to be introduced by Facebook called "Your Time on Facebook," which allows for management of time on Facebook

and includes an activity dashboard, a daily reminder, and management of notifications [19].

Methods

Participants

The participants of the survey were 200 postgraduates from a postgraduate support Facebook group called Doctorate Support Group (DSG). DSG is a support group that aims to provide a platform in supporting postgraduates to exchange ideas, expertise, and experiences in pursing their postgraduate education. The community has over 14,700 users who consist of postgraduates carrying out their studies and ex-postgraduate students who have completed their studies. The Facebook group applied a *community of practice* approach, in which the Facebook group provides a *shared knowledge* bank developed by the community members who have been associated with the community for a longer period. As newcomers join the group, the newcomers would learn from the *old-timers* who serve as *coaches* or *mentors* for the community. By participating in new activities and contributing to the community, the newcomers develop a new mastery of understanding and thus become recognized members (and later coaches) to give back to the community [3,20].

Data Collection and Analysis Approaches

The data were collected via administration of online surveys on Facebook group of the DSG. The online survey included mock interfaces (ie, high-fidelity prototypes) of the intervention features. This includes Facebook features as reported by Facebook in their upcoming tool called "Your Time on Facebook" [19]. These upcoming features allow for Facebook users to manage their time on Facebook, which includes an activity dashboard, a daily reminder, and management of notifications. As such, in investigating the possible Web-based intervention features for Facebook addiction during postgraduate studies, mobile phone addiction intervention features by Lee et al were adapted, which included manual monitoring, manual limit, automatic notification, automatic limit, and automatic reward features as shown in Table 1 [21]. In prevention of excessive Facebook usage, 2 features were assessed, which were manual monitoring and manual control. The manual monitoring feature investigated whether postgraduates perceived features such as manually monitoring their usage patterns (including usage time, frequency, locations, and mood) could help in Facebook addiction management. Meanwhile, the manual limit investigated whether features such as manual limiting Facebook usage based on time, location, Facebook features, and moods could potentially assist them in intervention of Facebook addiction. With regard to automatic notification, the Web-based features touched on interventions with regard to automatic notification of excessive Facebook usage based on location, features, time, and mood. The automatic limit feature is linked to features that automatically limit the usage based on time, location, frequency of use, and mood. The final feature (automatic reward), which is related to Web-based intervention features, is related to rewarding mechanisms based on usage duration and frequency, location, and mood.

The study also investigated aspects of Facebook addiction during postgraduates' studies using the Bergen Facebook Addiction Scale, which categorizes Facebook addiction disorder by psychological factors such as salience, tolerance, mood modification, relapse, withdrawal, and conflict [8]. As discussed before, salience is related to the mental state of continuously thinking about Facebook, whereas tolerance is related to tolerance level of Facebook usage. Mood modification is associated with whether Facebook affects current moods of the user, and relapse is linked with failed attempts of Facebook usage reduction, whereas withdrawal and conflict are related to negative conditions and effects because of failure in accessing Facebook [2,9]. Basic demographical data (ie, gender, age, and device usage on Facebook; current experience using Facebook; and Facebook usage frequency) were also assessed in the study.

The questionnaire developed based on Bergen Facebook Addiction Scale and Web-based intervention features (manual monitoring, manual limit, automatic notification, automatic limit, and automatic reward) was run through a content validation procedure to increase its level of validity. It was validated by 2 social networking analysis experts, 2 information technology experts, and a language lecturer. The content validation involved validation on aspects such as subject matter, technology, language, and measurement. This was conducted to verify that the questions for each variable were clear and concise. As a result, the online survey consisted of 48 items as measurements. The ethics committee of the Universiti Kebangsaan Malaysia approved the implementation of this study. We followed all national regulations and laws regarding human subjects' research and obtained the required permission to conduct this study. Participants provided Web-based informed consent to participate.

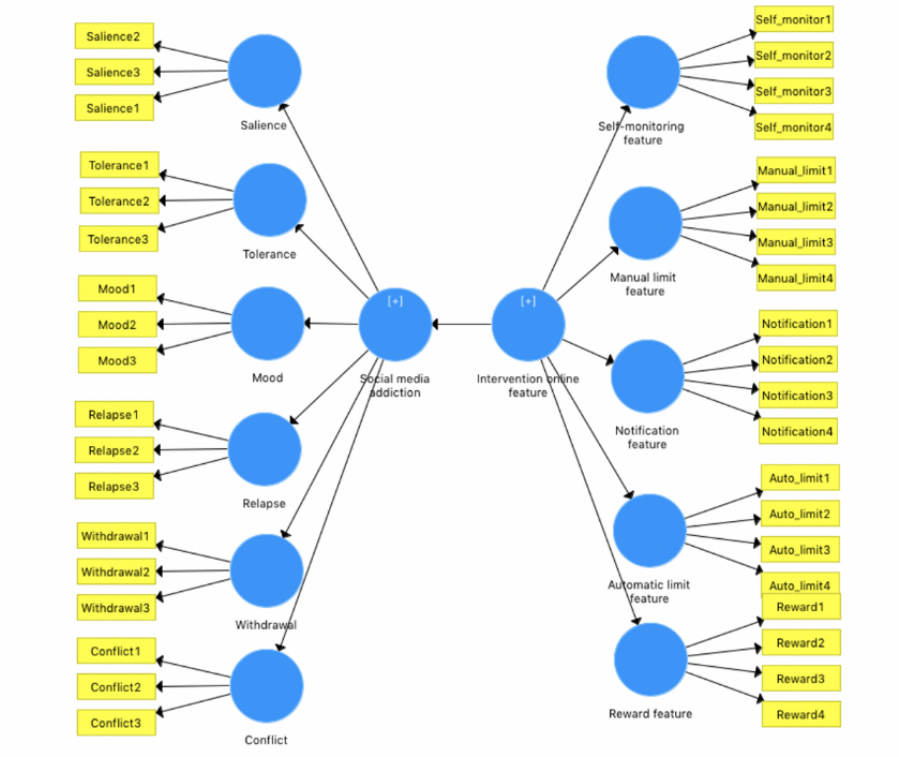
The data collected were analyzed using partial least square-structural equation modeling (PLS-SEM). This allowed for exploratory investigation of the relationship between Web-based intervention features for Facebook addiction and aspects of Facebook addiction [22]. As the study implements exploratory research in nature, PLS-SEM was chosen. The technique allows for conducting predictions and explanation of target constructs rather than confirmatory analysis with the capability of small samples sizes and complex models. PLS-SEM analysis does not make any assumptions about underlying data; thus, 200 Web-based participants are enough [21,22]. Hence, the study focuses only on describing the structural model analysis results (via PLS-SEM model diagrams) with regard to loadings of each construct and does not report on measurement model analysis results. The results of loadings would help in understanding the most and least important intervention features for Facebook addiction and the strongest and weakest causes of Facebook addiction among postgraduates. The software used to run PLS-SEM analysis is SmartPLS version 3.2.7 by SmartPLS GmbH. The model used in the analysis is illustrated in Figure 1.

Table 1. The constructs and respective indicators of the Web-based intervention features for Facebook addiction disorder.

Construct	Indicator
Self-monitoring feature (IF[a]_Self-monitoring)	• Track Facebook usage time • Track Facebook frequently used features • Track location of Facebook usage • Track mood while using Facebook
Manual limit feature (IF_Manual limit)	• Manually limit Facebook usage based on time • Manually limit Facebook usage based on location • Manually limit Facebook usage based on features • Manually limit Facebook usage based on mood
Notification feature (IF_Notifcation)	• Notification of excessive Facebook usage based on time • Notification of excessive Facebook usage based on location • Notification of excessive Facebook usage based on features • Notification of excessive Facebook usage based on mood
Automatic limit (IF_Auto-limit)	• Automatically limit Facebook usage based on time • Automatically limit Facebook usage based on location • Automatically limit Facebook features based on frequency of use • Automatically limit Facebook usage based on my mood
Reward feature (IF_Reward)	• Provide reward based on Facebook usage time • Provide reward based on Facebook usage location • Provide reward based on Facebook feature frequency • Provide reward based on mood using Facebook

[a]IF: intervention Web-based feature.

Figure 1. Model used for partial least square-structural equation modeling analysis for investigating the Web-based intervention features of Facebook addiction disorder.



Results

Demographical Findings

Most of the 200 participants of the online survey were female (77.5%, 155/200), whereas the remaining 22.5%, 45/200 were male. The age range was quite diverse as it included individuals aged between 23 years and 51 years. For devices used to access the Facebook, 9%, 18/200 of them used only computers, 36.5%, 73/200 used only mobile phones, and 53.5%, 107/200 of them used both computers and mobile phones. The survey also gained data regarding current experience of Facebook usage. The data revealed that 86.5%, 173/200 of them have used Facebook for more than 4 years, 11.5%, 23/200 of them have used Facebook for 3 to 4 years, whereas 2.0%, 4/200 of them have used Facebook for 1 to 2 years. In addition, Facebook usage frequency was also obtained. Findings indicated that most of them (83.0%, 166/200) access Facebook every day, whereas the others access it either 2 to 3 times a week or 4 to 5 times a week. Findings also revealed that 21.5%, 43/200 of them access Facebook more than 10 times a day, 20.0%, 40/200 access it 7 to 10 times a day, 24.0%, 48/200 access it 4 to 7 times a day, and 34.5%, 69/200 of them access it once a day.

Results on Partial Least Square-Structural Equational Modeling Results: Facebook Addiction

The findings of the measurement model analysis showed that the indicators of the constructs achieved internal consistency reliability, convergent validity, and divergent validity. The results are summarized in Tables 2-4. The findings of the structural measurement model analysis revealed that there are 6 addiction factors that are related to Facebook addiction in postgraduate studies, as shown in Figure 2. They are mood modification, withdrawal, tolerance, salience, conflict, and relapse, which received loadings of 0.5 and above [22]. This indicates that all the indicators (eg, FB_Salience1) are related to their respective constructs (eg, FB_Salience). These results corroborate with the works of Griffiths, Kuss and Griffiths, and Andreassen, where the studies revealed that these 5 levels contribute to SNS addiction, in this case, Facebook addiction [9,22,23].

Table 2. Internal consistency reliability results.

Indicator	Average variance extracted (AVE)	Composite reliability (CR)	R^2	Cronbach alpha
FB[a]_Conflict	0.732	0.891	0.593	.815
FB_Mood modification	0.722	0.886	0.394	.808
FB_Relapse	0.725	0.886	0.666	.804
FB_Salience	0.590	0.812	0.559	.659
FB_Tolerance	0.739	0.894	0.542	.821
FB_Withdrawal	0.845	0.942	0.511	.908
IF[b]_Auto-control	0.740	0.919	0.721	.882
IF_Manual Control	0.583	0.847	0.664	.759
IF_Notification	0.692	0.900	0.797	.852
IF_Reward	0.786	0.936	0.668	.909
IF_Self-monitoring	0.630	0.872	0.599	.803

[a]FB: Facebook addiction factor.

[b]IF: intervention Web-based feature.

Table 3. Convergent validity results.

Construct and indicator	Loading	Average variance extracted (AVE)	Composite reliability (CR)
FB[a]_Conflict		**0.732**	**0.891**
FB_Conflict1	0.845		
FB_Conflict2	0.923		
FB_Conflict3	0.794		
FB_Mood modification		**0.722**	**0.886**
FB_MoodModification1	0.837		
FB_MoodModification2	0.865		
FB_MoodModification3	0.847		
FB_Relapse		**0.725**	**0.886**
FB_Relapse1	0.709		
FB_Relapse2	0.910		
FB_Relapse3	0.918		
FB_Salience		**0.590**	**0.812**
FB_Salience1	0.705		
FB_Salience2	0.804		
FB_Salience3	0.792		
FB_Tolerance		**0.739**	**0.894**
FB_Tolerance1	0.765		
FB_Tolerance2	0.922		
FB_Tolerance3	0.884		
FB_Withdrawal		**0.845**	**0.942**
FB_Withdrawal1	0.928		
FB_Withdrawal2	0.948		
FB_Withdrawal3	0.880		
IF[b]_Auto-control		**0.740**	**0.919**
IF_AutoControl1	0.807		
IF_AutoControl2	0.876		
IF_AutoControl3	0.909		
IF_AutoControl4	0.845		
IF_Manual Control		**0.583**	**0.847**
IF_ManualControl1	0.815		
IF_ManualControl2	0.790		
IF_ManualControl3	0.769		
IF_ManualControl4	0.670		
IF_Notification		**0.692**	**0.900**
IF_Notification1	0.809		
IF_Notification2	0.858		
IF_Notification3	0.826		
IF_Notification4	0.834		
IF_Reward		**0.786**	**0.936**
IF_Reward1	0.856		

Construct and indicator	Loading	Average variance extracted (AVE)	Composite reliability (CR)
IF_Reward2	0.901		
IF_Reward3	0.918		
IF_Reward4	0.869		
IF_Self-monitor		**0.630**	**0.872**
IF_SelfMonitor1	0.768		
IF_SelfMonitor2	0.821		
IF_SelfMonitor3	0.840		
IF_SelfMonitor4	0.741		

[a]FB: Facebook addiction factor.

[b]IF: intervention Web-based feature.

Table 4. Divergent validity results (Fornell-Larcker Criterion).

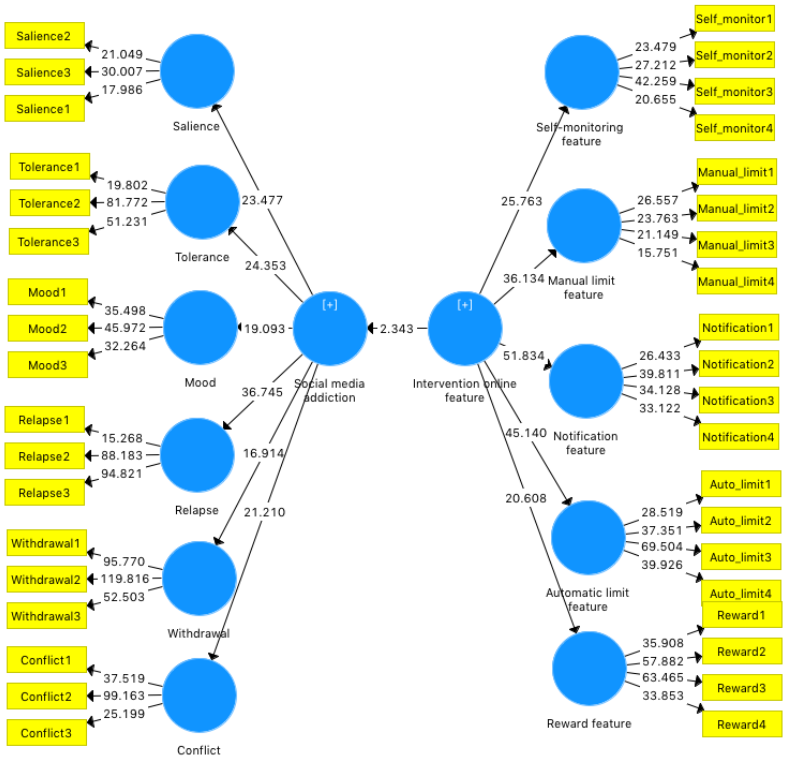
Construct	FB[a]_Conflict	FB_Mood modification	FB_Relapse	FB_Salience	FB_Tolerance	FB_Withdrawal	IF[b]_Auto-control	IF_Manual Control	IF_Notifi-cation	IF_Reward	IF_Self-monitoring
FB_Conflict	0.856	—[c]	—	—	—	—	—	—	—	—	—
FB_Mood modification	0.320	0.850	—	—	—	—	—	—	—	—	—
FB_Relapse	0.684	0.359	0.851	—	—	—	—	—	—	—	—
FB_Salience	0.483	0.406	0.505	0.768	—	—	—	—	—	—	—
FB_Tolerance	0.429	0.452	0.485	0.529	0.860	—	—	—	—	—	—
FB_Withdrawal	0.428	0.344	0.497	0.464	0.371	0.919	—	—	—	—	—
IF_Auto-control	0.160	0.251	0.226	0.106	0.209	0.038	0.860	—	—	—	—
IF_Manual Control	0.157	0.153	0.217	0.079	0.177	–0.015	0.626	0.763	—	—	—
IF_Notification	0.224	0.270	0.229	0.080	0.141	–0.005	0.718	0.697	0.832	—	—
IF_Reward	0.200	0.241	0.174	0.143	0.154	0.002	0.599	0.521	0.668	0.886	—
IF_Self-monitoring	0.115	0.162	0.166	0.141	0.140	0.012	0.539	0.610	0.597	0.545	0.794

[a]FB: Facebook addiction factor.

[b]IF: intervention Web-based feature.

[c]Not applicable.

Figure 2. Bootstrapping results of the partial least square-structural equation modeling analysis.



The results also indicated that the Facebook addiction construct with the highest loading was relapse (0.666), followed by conflict (0.593), salience (0.559), tolerance (0.542), and withdrawal (0.511), as in Table 5. The lowest loading was obtained by the mood modification construct that was 0.394. The results signify that the strongest Facebook addiction factors in postgraduates' studies are relapse and conflict, whereas the 2 weakest levels are mood modification and withdrawal. This contradicts the results of the studies by Koc and Gulyagci as well as Balakrishnan and Shamim, where the former authors revealed that mood modification and conflict are the most frequent symptoms of Facebook addictive usage among university students, whereas the latter authors revealed that salience, loss of control, and withdrawal are the main indicators of Facebook addiction among students [24,25].

Table 5. Structural model results.

Hypothesis	R^2	Beta	SE	*t* value	Decision
FB[a]_Conflict ≥ FB Addiction	0.593	.238	0.020	12.144[b]	Support
FB_Mood modification ≥ FB Addiction	0.394	.190	0.022	8.680[b]	Support
FB_Relapse ≥ FB Addiction	0.666	.251	0.018	13.929[b]	Support
FB_Salience ≥ FB Addiction	0.559	.184	0.017	10.672[b]	Support
FB_Tolerance ≥ FB Addiction	0.542	.229	0.019	12.384[b]	Support
FB_Withdrawal ≥ FB Addiction	0.511	.256	0.020	12.966[b]	Support
Intervention Features ≥ IF[c]_Auto-control	0.721	.851	0.024	35.550[b]	Support
Intervention Features ≥ IF_Manual Control	0.664	.816	0.036	22.926[b]	Support
Intervention Features ≥ IF_Notification	0.797	.891	0.025	36.247[b]	Support
Intervention Features ≥ IF_Reward	0.668	.820	0.037	22.182[b]	Support
Intervention Features ≥ IF_Self-monitoring	0.599	.776	0.045	17.367[b]	Support

[a]FB: Facebook addiction factor.

[b]$P<.05$.

[c]IF: intervention Web-based feature.

These results could be caused by the fact that Facebook addiction factors could potentially be explained by a process in which a Facebook addict goes through levels of addictions that ends with relapse and conflict levels, where they attempt to reduce Facebook time but fail to do so (relapse) and ignore their studies and people (conflict) [9]. This can also be related to Facebook usage frequency of postgraduates in this study, where most of them (83%) accessed Facebook every day and 65.5% of them accessed Facebook more than 4 times a day. In addition, students who are Facebook addicts have possibly never deactivated their accounts before showing their high Facebook addiction level [25]. Furthermore, Cabral reported that the majority of participants in their study reported failed attempts of social media usage reduction [26].

The findings also revealed that 2 of the relapse construct's indicators FB_Relapse2 and FB_Relapse2 obtained the highest loadings. The indicators were related to relapse in decision making and actions on Facebook usage, which included "decided to use Facebook during your postgraduate studies less frequently, but not managed to do so" and "tried to cut down on the use of Facebook during your postgraduate studies without success." This is in line with the findings of Brailovskaia and Margraf's study that investigated Facebook addiction disorder among German students [2]. They discovered that Facebook addiction factors fully mediated the association between narcissism and stress systems, and the highest positive association was with 3 factors, which were relapse, withdrawal, and salience. From that study, they revealed that users who are narcissist tend to spend more time thinking about Facebook because of Web-based self-presentation, interaction, and reflections in the social networking platform, thus causing them to be vulnerable to Facebook addiction and be in a state of relapse.

Discussion

Discussion on Partial Least Square-Structural Equational Modeling Results: Web-Based Intervention Features

The findings of the structural measurement model analysis show that 6 Web-based intervention features are related to Web-based intervention and Facebook addiction in postgraduate studies. The factors are manual monitoring feature, manual limit feature, automatic notification feature, automatic limit feature, and automatic reward feature, which obtained loadings above the 0.5 cut-off point for loadings, as shown in Tables 2-4 [22]. This indicates that all the indicators (eg, IF_manual_monitoring1) are related to their respective constructs (eg, manual monitoring).

The results also revealed that the Web-based intervention feature for postgraduate education that received the highest loading was automatic notification feature (0.797), followed by automatic limitation feature (0.721), automatic reward feature (0.668), and manual limitation feature (0.664). The lowest loading gained was by manual monitoring feature (0.599), as shown in Table 6. The results suggest that the 5 intervention features could be used in management or intervention of Facebook addiction in postgraduate education. In other words, this indicates that postgraduates prefer to be notified of their Facebook usage (notification) and then be automatically managed or restricted to Facebook based on time, frequency, and location of Facebook usage as well as mood during Facebook access. Although this may seem like a straightforward solution in managing Facebook addiction, it may not be the case. This can be related to a study on Facebook addiction with regard to active Facebook use (ie, using Facebook for communication) and passive Facebook use (ie, using Facebook to consume content) [5]. They discovered that passive Facebook use was related to daily life events. Interestingly, the study revealed that participants of the study increased Facebook usage following positive life events instead of negative ones. In other words, passive Facebook use is less likely to be associated with escapism as users have decreased level of passive Facebook use when faced with problems as compared with positive experiences.

This can further be related to another relevant study, where the study indicated that Facebook addiction is related to narcissism and stress systems [2]. Linking the 2 studies together, this indicates that postgraduates who have Facebook addicts are more likely to use Facebook to consume information-related positive life events, in this case related to academic success, rather than using Facebook for escapism related to negative emotions. On that note, it would be interesting for future Web-based intervention features to include the option for passive and active Facebook use and relate it with positive and negative life events in postgraduate education. In terms of the automatic reward feature, this suggests that rewards (eg, rewards systems in gaming, such as scores, or virtual currencies—refer to Yen's study [27]) could be used as an intervention measure for addicts. Although results revealed that manual control and self-monitoring were the least important intervention features, both are still essential as they allow postgraduates to monitor their Facebook usage levels and manually control/manage Facebook features based on time, location, and feature usage as well as inputting their moods.

Table 6. Coefficient of determination (R^2) test.

Hypothesis	R^2
FB[a]_Relapse ≥ FB Addiction	0.666
FB_Conflict ≥ FB Addiction	0.593
FB_Salience ≥ FB Addiction	0.559
FB_Tolerance ≥ FB Addiction	0.542
FB_Withdrawal ≥ FB Addiction	0.511
FB_Mood modification ≥ FB Addiction	0.394
Intervention Features ≥ IF[b]_Notification	0.797
Intervention Features ≥ IF_Auto-control	0.721
Intervention Features ≥ IF_Reward	0.668
Intervention Features ≥ IF_Manual Control	0.664
Intervention Features ≥ IF_Self-monitoring	0.599

[a]FB: Facebook addiction factor.

[b]IF: intervention Web-based feature.

Conclusions, Implications, and Future Directions

The study discovered 6 Facebook addiction factors (relapse, conflict, salience, tolerance, withdrawal, and mood modification) and 5 intervention features (notification, auto-control, reward, manual control, and self-monitoring) that could be used in management of Facebook addiction in postgraduate education. The study also revealed that relapse is the most important factor and mood modification is the least important factor. Furthermore, findings indicated that notification was the most important intervention feature, whereas self-monitoring was the least important feature. This study's findings, with regards to social media addiction factors and Web-based intervention features, could assist future developed and educators in the development of Web-based intervention tools for Facebook addiction management in postgraduate education. In addition, PLS-SEM was used as a statistical approach to verify the relationship between social media addiction disorder management and Web-based intervention features, which contributes to the field in terms of the higher education field, particularly in postgraduate education.

Future directions in this area are as follows. First, the addiction factors and intervention features were only tested in postgraduate educational settings. It would be interesting to investigate whether the findings corroborate or contradict with these findings in other educational settings, which include undergraduate, primary, and secondary education as well as long-life learning settings [28]. Second, most of the respondents were studying in local higher education institutions. It would be worth replicating the study with a larger sample with a more diverse span of international higher educational institutions [29,30]. Finally, it would be interesting to combine the results with social network analysis as to indicate whether social network patterns (in egocentric diagrams) could be used in the management of Facebook addiction [3] as well as other Web-based approaches [31-36].

Acknowledgments

This research is funded by the European Union Erasmus Mundus Action 2 Techno II Project, the Malaysian Research Universities Network (MRUN) Translational Program Grant (Grant No: MRUN-RAKAN RU-2019-003/2), and FPEND Research Grant (Grant No: GG-2019-064).

Conflicts of Interest

None declared.

References

1. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders (DSM-5®). Washington, DC: American Psychiatric Pub; 2013.
2. Brailovskaia J, Margraf J. Facebook addiction disorder (FAD) among German students-a longitudinal approach. PLoS One 2017;12(12):e0189719 [FREE Full text] [doi: 10.1371/journal.pone.0189719] [Medline: 29240823]
3. Ryan T, Chester A, Reece J, Xenos S. The uses and abuses of Facebook: a review of Facebook addiction. J Behav Addict 2014 Sep;3(3):133-148 [FREE Full text] [doi: 10.1556/JBA.3.2014.016] [Medline: 25317337]

4. Stavropoulos V, Kuss DJ, Griffiths MD, Wilson P, Motti-Stefanidi F. MMORPG gaming and hostility predict internet addiction symptoms in adolescents: an empirical multilevel longitudinal study. Addict Behav 2017 Jan;64:294-300. [doi: 10.1016/j.addbeh.2015.09.001] [Medline: 26410795]
5. Young NL, Kuss DJ, Griffiths MD, Howard CJ. Passive Facebook use, Facebook addiction, and associations with escapism: an experimental vignette study. Comput Human Behav 2017 Jun 1;71:24-31. [doi: 10.1016/j.chb.2017.01.039]
6. Andreassen CS, Pallesen S. Social network site addiction - an overview. Curr Pharm Des 2014;20(25):4053-4061. [doi: 10.2174/13816128113199990616] [Medline: 24001298]
7. Mazzoni E, Cannata D, Baiocco L. Focused, not lost: the mediating role of temporal dissociation and focused immersion on problematic internet use. Behav Inform Technol 2016 May 17;36(1):1-10. [doi: 10.1080/0144929X.2016.1159249]
8. Andreassen CS, Torsheim T, Brunborg GS, Pallesen S. Development of a Facebook addiction scale. Psychol Rep 2012 Apr;110(2):501-517. [doi: 10.2466/02.09.18.PR0.110.2.501-517] [Medline: 22662404]
9. Andreassen CS. Online social network site addiction: a comprehensive review. Curr Addict Rep 2015 Apr 11;2(2):175-184. [doi: 10.1007/s40429-015-0056-9]
10. Hong FY, Huang DH, Lin HY, Chiu SL. Analysis of the psychological traits, Facebook usage, and Facebook addiction model of Taiwanese university students. Telemat Inform 2014 Nov 1;31(4):597-606. [doi: 10.1016/j.tele.2014.01.001]
11. Sofiah SZ, Omar SZ, Bolong J, Osman MN. Facebook addiction among female university students. Revista de Administratie Publica si Politici Sociale 2011 Dec 1;3(7):95-109.
12. Sheldon P. The relationship between unwillingness-to-communicate and students' Facebook use. J Media Psychol 2008 Jan;20(2):67-75. [doi: 10.1027/1864-1105.20.2.67]
13. Griffiths MD. Social networking addiction: emerging themes and issues. J Addict Res Ther 2013;4(5):1000e118. [doi: 10.4172/2155-6105.1000e118]
14. Wolniczak I, Cáceres-DelAguila JA, Palma-Ardiles G, Arroyo KJ, Solís-Visscher R, Paredes-Yauri S, et al. Association between Facebook dependence and poor sleep quality: a study in a sample of undergraduate students in Peru. PLoS One 2013;8(3):e59087 [FREE Full text] [doi: 10.1371/journal.pone.0059087] [Medline: 23554978]
15. Karaiskos D, Tzavellas E, Balta G, Paparrigopoulos T. P02-232 - social network addiction: a new clinical disorder? Eur Psychiatry 2010 Jan 1;25(1):855. [doi: 10.1016/S0924-9338(10)70846-4]
16. Çam E, Isbulan O. A new addiction for teacher candidates: social networks. Turk Online J Educ Technol 2012 Jul;11(3):14-19 [FREE Full text]
17. Atroszko PA, Balcerowska JM, Bereznowski P, Biernatowska A, Pallesen S, Andreassen CS. Facebook addiction among Polish undergraduate students: validity of measurement and relationship with personality and well-being. Comput Human Behav 2018 Aug;85:329-338. [doi: 10.1016/j.chb.2018.04.001]
18. Turel O, Brevers D, Bechara A. Time distortion when users at-risk for social media addiction engage in non-social media tasks. J Psychiatr Res 2018 Feb;97:84-88. [doi: 10.1016/j.jpsychires.2017.11.014] [Medline: 29220826]
19. Jafarkarimi H, Sim AT, Saadatdoost R, Hee JM. Facebook addiction among Malaysian students. Int J Innov Eng Technol 2016 Jun 1;6(6):465-469. [doi: 10.7763/IJIET.2016.V6.733]
20. Facebook Newsroom. 2018. New Tools to Manage Your Time on Facebook and Instagram URL: https://newsroom.fb.com/news/2018/08/manage-your-time/ [accessed 2018-08-29]
21. Wenger E, White N, Smith JD. Digital Habitats: Stewarding Technology for Communities. Portland, Oregon: CPsquare; 2009.
22. Norman H, Nordin N, Din R, Ally M, Dogan H. Exploring the roles of social participation in mobile social media learning: a social network analysis. Int Rev Res Open Distr Learn 2015 Nov 2;16(4):205-224. [doi: 10.19173/irrodl.v16i4.2124]
23. Lee H, Ahn H, Choi S, Choi W. The SAMS: smartphone addiction management system and verification. J Med Syst 2014 Jan;38(1):1. [doi: 10.1007/s10916-013-0001-1] [Medline: 24395031]
24. Hair Jr J, Hult GT, Ringle CM, Starstedt M. A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM). Thousand Oaks, CA: Sage Publications; 2016.
25. Griffiths MA. A 'components' model of addiction within a biopsychosocial framework. J Subst Use 2009 Jul 12;10(4):191-197. [doi: 10.1080/14659890500114359]
26. Koc M, Gulyagci S. Facebook addiction among Turkish college students: the role of psychological health, demographic, and usage characteristics. Cyberpsychol Behav Soc Netw 2013 Apr;16(4):279-284. [doi: 10.1089/cyber.2012.0249] [Medline: 23286695]
27. Balakrishnan V, Shamim A. Malaysian Facebookers: motives and addictive behaviours unraveled. Comput Human Behav 2013 Jul 1;29(4):1342-1349. [doi: 10.1016/j.chb.2013.01.010]
28. Cabral J. Is generation y addicted to Facebook. J Undergrad Res Commun 2011;2(1):5-14 [FREE Full text]
29. Yen JY, Yen CF, Chen CS, Wang PW, Chang YH, Ko CH. Social anxiety in online and real-life interaction and their associated factors. Cyberpsychol Behav Soc Netw 2012 Jan;15(1):7-12 [FREE Full text] [doi: 10.1089/cyber.2011.0015] [Medline: 22175853]
30. Ashford RD, Lynch K, Curtis B. Technology and social media use among patients enrolled in outpatient addiction treatment programs: cross-sectional survey study. J Med Internet Res 2018 Mar 6;20(3):e84 [FREE Full text] [doi: 10.2196/jmir.9172] [Medline: 29510968]

31. Khalili-Mahani N, Smyrnova A, Kakinami L. To each stress its own screen: a cross-sectional survey of the patterns of stress and various screen uses in relation to self-admitted screen addiction. J Med Internet Res 2019 Apr 2;21(4):e11485 [FREE Full text] [doi: 10.2196/11485] [Medline: 30938685]
32. Faraci P, Craparo G, Messina R, Severino S. Internet Addiction Test (IAT): which is the best factorial solution? J Med Internet Res 2013 Oct 9;15(10):e225 [FREE Full text] [doi: 10.2196/jmir.2935] [Medline: 24184961]
33. Lai MH, Maniam T, Chan LF, Ravindran AV. Caught in the web: a review of web-based suicide prevention. J Med Internet Res 2014 Jan 28;16(1):e30 [FREE Full text] [doi: 10.2196/jmir.2973] [Medline: 24472876]
34. Jaafar NR, Bahar N, Ibrahim N, Ismail WS, Baharudin A. Excessive internet use in young women: what are the implications? Curr Opin Psychiatry 2017 Jul;30(4):260-267. [doi: 10.1097/YCO.0000000000000336] [Medline: 28426547]
35. Ali NM, Shahar S, Kee YL, Norizan AR, Noah SA. Design of an interactive digital nutritional education package for elderly people. Inform Health Soc Care 2012 Dec;37(4):217-229. [doi: 10.3109/17538157.2012.654843] [Medline: 22583111]
36. Bergin P, Sadleir L, Legros B, Mogal Z, Tripathi M, Dang N, EpiNet Study Group. An international pilot study of an internet-based platform to facilitate clinical research in epilepsy: the EpiNet project. Epilepsia 2012 Oct;53(10):1829-1835 [FREE Full text] [doi: 10.1111/j.1528-1167.2012.03636.x] [Medline: 22958055]

Abbreviations

DSG: Doctorate Support Group
FB: Facebook addiction factor
IF: intervention Web-based feature
PLS-SEM: partial least square-structural equation modeling
SNSs: social networking sites

Edited by G Eysenbach; submitted 27.05.19; peer-reviewed by EK Özyirmidokuz, GE Iyawa; comments to author 29.06.19; revised version received 13.07.19; accepted 19.07.19; published 02.10.19.

Please cite as:
Dogan H, Norman H, Alrobai A, Jiang N, Nordin N, Adnan A
A Web-Based Intervention for Social Media Addiction Disorder Management in Higher Education: Quantitative Survey Study
J Med Internet Res 2019;21(10):e14834
URL: https://www.jmir.org/2019/10/e14834
doi:10.2196/14834
PMID:31579018

Original Paper

Improving the Course of Depressive Symptoms After Inpatient Psychotherapy Using Adjunct Web-Based Self-Help: Follow-Up Results of a Randomized Controlled Trial

Rüdiger Zwerenz[1], PhD; Carlotta Baumgarten[1], MSc; Jan Becker[2], PhD; Ana Tibubos[1], PhD; Martin Siepmann[3,4], MD; Rudolf J Knickenberg[3], MD; Manfred E Beutel[1], MD

[1]Department of Psychosomatic Medicine and Psychotherapy, University Medical Center, Johannes Gutenberg-University, Mainz, Germany

[2]Institute for Teachers' Health, University Medical Center, Johannes Gutenberg-University, Mainz, Germany

[3]Clinic for Psychosomatic Medicine, Rhön-Klinikum AG, Bad Neustadt/ Saale, Germany

[4]Department of Psychotherapy and Psychosomatic Medicine, Faculty of Medicine Carl Gustav Carus, Technische Universität Dresden, Dresden, Germany

Corresponding Author:
Rüdiger Zwerenz, PhD
Department of Psychosomatic Medicine and Psychotherapy
University Medical Center
Johannes Gutenberg-University
Untere Zahlbacher Str 8
Mainz, 55131
Germany
Phone: 49 61311 75981
Fax: 49 6131 17475981
Email: ruediger.zwerenz@unimedizin-mainz.de

Abstract

Background: We recently showed in a randomized controlled trial that Web-based self-help as an adjunct improved the effectiveness of multimodal inpatient psychotherapy for depression.

Objective: The aims of this study were (1) to determine whether a Web-based self-help adjunctive to multimodal inpatient psychotherapeutic treatment could also improve the course of depressive symptoms and (2) to identify predictors of residual depressive symptoms at follow-up.

Methods: Overall, 229 patients were randomized either to the Web-based self-help intervention group (Deprexis) or an active control group (Web-based information about depression and depressive symptoms) in addition to multimodal inpatient psychotherapy. Participants in both groups were able to access their respective Web-based programs for 12 weeks, which meant that they typically had access after discharge from the inpatient unit (mean hospitalization duration: 40 days, T1). Follow-up was performed 6 months after study intake (T3).

Results: At follow-up, participants of the Web-based self-help group had considerably lower symptom load regarding depressive symptoms (d=0.58) and anxiety (d=0.46) as well as a better quality of life (d=0.43) and self-esteem (d=0.31) than participants of the control group. Nearly 3 times as many participants of the intervention group compared with the control group achieved remission in accordance with less deterioration. The number needed to treat based on the Beck Depression Inventory-II (BDI-II) improved over time (T1: 7.84, T2: 7.09, and T3: 5.12). Significant outcome predictors were BDI at discharge and treatment group.

Conclusions: Web-based self-help as an add-on to multimodal inpatient psychotherapy improved the short-term course of depressive symptoms beyond termination. Residual symptoms at discharge from inpatient treatment and utilization of the Web-based self-help were the major predictors of depressive symptoms at follow-up. Challenges and barriers (eg, costs, therapists' concerns, or technical barriers) of adding Web-based interventions to inpatient treatment have to be addressed.

Trial Registration: ClinicalTrials.gov NCT02196896; https://clinicaltrials.gov/ct2/show/NCT02196896.

(J Med Internet Res 2019;21(10):e13655) doi:10.2196/13655

KEYWORDS

depression; mental health; internet; aftercare; psychotherapy; psychology, clinical; inpatients

Introduction

Benefit and Effectiveness of Web-Based Interventions

Depression has been recognized as one of the leading health problems with a 12-month prevalence of 6.9% [1] and an even higher prevalence of depression and depressive symptoms in outpatients of different clinical specialties (27%) [2]. Given its high prevalence and limited access to evidence-based treatments [3], Web-based self-help interventions have been developed to provide instant, flexible help for a great variety of mental health problems [4]. Several meta-analyses have shown that they are effective treatments of depressive symptoms with effect sizes comparable with face-to-face treatments [5] and have been proven to prevent successfully the recurrence in remitted depression patients [6,7]. Efficacy, however, depends on contextual factors Web-based interventions are implemented, such as a meta-analysis [8] of the Web-based intervention moodgym showed. Sensitivity analyses have shown that general efficacy (g=0.36) is lower (g=0.17), when publication-bias is considered, an active control group is used as a comparative condition (g=0.12 vs g=0.53), guidance is missing (g=0.23 vs g=0.75), or adherence is low (g=0.22 vs g=0.64); on the other hand efficacy is higher in some countries, such as Australia (g=0.78) than in others (eg, Europe g=0.17).

Deprexis is an interactive Web-based self-help program [9] with proven effectiveness in several randomized controlled studies (RCTs). In a recent meta-analysis of self-guided Web-based interventions for the treatment of depressive symptoms [10], 13 trials of which 5 were Deprexis trials have been combined, and the overall effect was g=0.31. Compared with the meta-analysis of Karyotaki [10] with heterogenous internet-based cognitive behavioral therapy interventions, a recent meta-analysis based on 8 studies with Deprexis reported a slightly higher posttreatment between-group effect size for the improvement of depressive symptoms of g=0.54 (95% CI 0.39-0.69) [11]. Moreover, there is evidence for efficacy after completing the intervention with between-group effect size of d=0.32 in a 6-month follow-up compared with usual psychological or pharmacological treatments alone [12]. In another study, between-group effect sizes from d=0.36 at 3 months to d=0.13 at 12 months were reported, and remission rates over time were significantly higher in the group receiving Deprexis [13].

Web-Based Interventions in Regular Care and Blended Treatments

However, the challenge has remained how to integrate Web-based treatments into regular care [5], especially as effectiveness studies show quite mixed results. In a large RCT of 2 Web-based cognitive behavioral interventions, which had proven to be effective in prior RCTs, no statistical significant effect could be shown when these interventions have been added to usual care [14]. In an RCT, adding an internet-based relapse prevention program to treatment as usual (TAU) was not cost-effective regarding depression-free days and quality adjusted life years compared with TAU alone [15].

In blended treatments, face-to-face interventions are combined with Web-based or mobile-based interventions and integrated in 1 treatment scheme [16]. Nakao et al [17] compared a Web-based program combined with face-to-face sessions with regular cognitive behavioral therapy in an RCT with N=40 participants and could show that blended CBT was effective in reducing depressive symptoms in patients with major depression. There are also first promising studies on blended group psychotherapy; in a qualitative study, patients reported in in-depth interviews that blended group therapy could be motivating, increasing the consolidation of the results of face-to-face psychotherapy, among other things, because of the possibility of between-session monitoring and Web-based reinforcement of exercising tasks [18]. In a randomized controlled feasibility study, large between-group effect sizes (d=0.87) could be shown in favor of blended group psychotherapy compared with a wait list control group [19]. However, the often-proven cost-effectiveness of certain Web-based interventions (eg, McCrone et al [20]) is not automatically the case for blended treatments. In a naturalistic study, 4448 records of patients with depressive symptoms or anxiety were compared using propensity score matching whether treatment was done in a regular face-to-face setting or blended treatment; 1 main result was that blended treatment resulted in significantly higher costs mainly because of more treatment time than in the regular face-to-face setting [21].

Web-Based Interventions as an Add-On to Inpatient Psychotherapy

Inpatient psychotherapy is indicated in severe, chronic and complex cases of depression, compounded by mental or somatic comorbid conditions according to German medical guidelines (DGPPN 2015) [22]. It has reduced depressive symptoms with a large effect size (d=1.2) based on the Beck Depression Inventory (BDI) after 61.8 days of inpatient treatment in a multicenter study [23], respectively; Hedges g=0.84 according to a meta-analysis [24]. On the basis of the Quick Inventory of Depressive Symptoms expert rating), Zeeck et al [25] reported complete remission in 29% of patients, partial remission in 11%, modest change of 31%, and nonresponse in 29% in a multicenter study with a mean inpatient or day hospital treatment duration of 10 weeks. However, even after successful inpatient treatment, there may be residual symptoms that increase the risk of relapse or recurrence of depressive symptoms [26,27].

In an RCT with 229 depression inpatients [28], we compared Deprexis with an active control group of psychoeducation comprising a weekly Web-based information regarding etiology and treatment of depression. Both conditions were adjuncts to intensive psychodynamic inpatient psychotherapy with a mean duration of 40 days (range 11-78 days). As participants were eligible to use Deprexis for a total of 12 weeks, they could continue to use it for an average of at least 6 weeks after discharge from treatment. Depressive symptoms were

significantly lower in the intervention group at discharge from inpatient treatment and at the end of intervention (3 months after study inclusion), with a moderate effect size (d=0.47 at discharge and d=0.44 3 months after study inclusion). This counted also for anxiety, quality of life, and self-esteem with effect sizes between d=0.33 and d=0.38 at discharge [28].

The aim of this study was to determine whether the add-on of a Web-based self-help program to multimodal inpatient treatment of depressed patients also improves stability of remission and test postulated predictors of residual depression [29] at follow-up based on patient characteristics. As residual symptoms were lower in the group with adjunct Web-based self-help versus psychoeducation, we hypothesized that remission rates would be higher at follow-up 6 months after study inclusion.

Methods

Participants

From July 2014 to February 2016, patients have been recruited in the Psychosomatic Clinic in Bad Neustadt/Saale, Germany. Eligible patients were aged between 18 to 65 years, had private internet access, sufficient German language skills, a score in the BDI-II above 13, and a clinical diagnosis of depression (International Statistical Classification of Diseases and Related Health Problems 10th revision: F32.x, F33.x, F34.1, and F43.2). Patients were excluded if a diagnosis of (1) psychosis (F20-F29); (2) current alcohol or drug addiction (F10-F19); (3) borderline (F60.3), antisocial (F60.2), schizoid (F60.1), and schizotypal (F21) personality disorders; (4) anorexia nervosa (F50.0); and (5) lifetime diagnoses of schizophrenia (F20-F29), schizoaffective (F25), bipolar (F31), or organic (F00-F09) mental disorder was present.

Eligible patients received oral and written information about the study and its requirements as part of a weekly information session. Therapists were introduced to the rationale of the intervention but had no active part in 1 of the add-on treatments.

After signing written informed consent, participants were coded and randomized to one of the groups (intervention group vs control group). Procedure and study protocol were conducted in accordance with the declaration of Helsinki and approved by the ethics committee of the Statutory Physician Board of the State of Rhineland Palatinate (Ref No 837.093.14 [9332-F]). The trial protocol was published elsewhere [30].

As described in more detail in the study by Zwerenz et al [28], 611 patients were eligible to participate in the study, 135 of whom did not meet inclusion criteria, 180 of whom did not want to participate, and 67 of whom did not complete study consent. Accordingly, out of the remaining 229 patients randomized, 215 patients were analyzed (N=108 in the intervention group and N=107 in the control group) who had received the respective intervention. At the end of inpatient treatment (T1), 198/229 (86.5%) participants completed the assessment: 87.8% (101/115) in the intervention group and 85.0% (97/114) in the control group. At the end of the intervention (T2) 74.2% (170/229) of the participants completed the assessment 73.9% (85/115) in the intervention group and 74.5% (85/114) in the control group. At the follow-up (T3), 69.9% (160/229) completed the assessment, 75.6% (87/115) in the intervention group and 64.0% (73/114) in the control group. Participants who completed the T2 and follow-up assessments did not differ from those participants who dropped out from assessments concerning baseline mental symptoms or self-assessed work ability.

Participants were predominantly female (60.7%; 139/229) and had a mean age of 48 years (SD 9.79), ranging from 18 to 65 years. About half of the participants were married (50.2%; 115/229), graduated from middle or higher secondary level (58.1%; 133/229), and worked full-time (47.6%; 109/229), with 56.3% (129/229) being on sick leave at study intake. Mean inpatient treatment duration was 40 days (range 11-78; SD 7.51), with no difference between intervention group (mean 41, SD 7.43) and control group (mean 40, SD 7.58). Previous psychopharmacological and psychotherapeutic treatments were comparable in the intervention group and control group as well as the status of antidepressant medication during inpatient treatment [28]. As delineated previously [28], the majority (79.9%; 183/229) reported having accessed the intervention or the Web-based material at least once. Furthermore, almost twice as many participants used Deprexis (46.0%; 53/115) regularly, that is, several times a week, compared with the utilization of the Web-based information in the control group (23.6%; 27/114).

Intervention and Control Condition

In the Psychosomatic Clinic Bad Neustadt, multimodal inpatient psychodynamic psychotherapy entails individual and group psychotherapy, creative psychotherapy interventions, and adjunct treatments such as patient education and exercising. Insight-oriented group therapy was combined with nonverbal treatment methods, that is, body therapy to address difficulties in emotion modulation, interpersonal problems, core beliefs about oneself, deficits in self-esteem, and self-care that contributed to the depressive symptoms. In addition to multimodal inpatient psychodynamic psychotherapy, participants of the intervention group got access to the Web-based self-help program Deprexis for 12 weeks. It consists of 10 main modules plus an introductory and a summary module based on cognitive behavioral techniques, positive psychology, emotion-focused therapy, and dream work (for details, cf. Meyer et al [9]). An interactional dialogue guides the user presenting text blocks with optional graphics, exercises, audio files, and answering options. Subsequent text blocks are based on the user's choices. A new module is presented only after completing the prior module. Optional reminders can be activated via email and short message service (see Meyer et al [9] and Zwerenz et al [28]). During inpatient treatment, participants had 2 1-hour time slots implemented in their weekly treatment plan, when they got access to a computer. Back home, continued access to Deprexis was provided until the period of 12 weeks had expired.

For participants of the control group, an internet platform was accessible in addition to the inpatient psychotherapy (TAU), consisting of 12 weekly modules with specific topics regarding depression, for example, information on depressive disorders, treatment options (psychotherapeutic and medication), efficacy of different treatments, and prognostic factors. This information

was mainly based on the official treatment guidelines for major depression in Germany [22]. Analogous to the intervention group, the treatment plan provided participants with time slots and computer access for using the Web-based platform [28,30]. All treatments were performed by the same group of therapists.

Outcome Measures

In addition to sociodemographic variables and relevant treatment data (ie, previous treatment, diagnosis, medication, treatment duration, etc), standardized questionnaires were used. Primary and secondary outcome measures were collected by self-report using the Web-based survey platform SoSci Survey (SoSci Survey GmbH [31]).

Depressive symptoms, our primary study outcome measured by the BDI-II [30], a reliable and valid instrument [32,33], were assessed at baseline (T0), discharge from the hospital (T1), termination of the program (T2), and follow-up (T3) 6 months after study admission. Apart from the time of discharge from the clinic (T1), secondary outcomes were also surveyed.

Secondary outcomes were assessed by well-established, reliable, and valid measures. Depressive symptoms were additionally measured with the Patient Health Questionnaire-9 (PHQ-9 [34]; Cronbach alpha=.85). Generalized anxiety was assessed with the generalized anxiety disorder-7 (GAD-7 [35]; Cronbach alpha=.92), quality of life (assessed with the European Health Interview Survey Quality of Life 8-Item index, EUROHIS-QOL 8-item index [36]; Cronbach alpha=.78), and self-esteem by the Rosenberg Self-esteem Scale (RSE [37]; Cronbach alpha=.84). Dysfunctional attitudes related to depressive thinking were measured with the Dysfunctional Attitudes Scale (DAS [38]; Cronbach alpha=.86) and work ability by the short version of the Work Ability Index (WAI [39]; Cronbach alpha=.80). Satisfaction, positive and negative influence, as well as adherence to the Web-based interventions were measured by single items on 5-point Likert scales. As potential predictor of outcome, childhood trauma was assessed with the German version of the Childhood Trauma Questionnaire [40] (Cronbach alpha=.89). To measure structural psychological deficits, the valid short form of the Operationalized Psychodynamic Diagnosis structured questionnaire [41] was used (Cronbach alpha=.61 and .87).

Data Analyses

All analyses were performed on the basis of intention to treat. To replace missing values, we used a Markov Chain Monte Carlo multivariate imputation algorithm with IBM SPSS Statistics 23 and 5 imputations, 10 estimations per missing value, and a constraint of a maximum of 60% missing data.

Outcome measures were evaluated in both conditions by analyses of covariance with baseline scores as covariate. Remission rates, significant reliable change [42], and number needed to treat (NNT) were calculated. As a reversal of the relative risk, the NNT indicates the number of patients that have to be treated to generate an additional positive outcome in 1 of them [43]. Comparisons of the between-group effects with regard to these variables were tested with χ^2 tests. To test the postulated predictors of depressive symptoms at follow-up [30], we performed multivariate analysis using a generalized linear model (GLM). According to protocol [30], depressive symptoms at baseline, childhood trauma, degree of structural deficits, and utilization of other (psychotherapeutic, pharmacological, inpatient, and self-help) treatments after the inpatient treatment were included as independent variables controlling for sex and age. Group (intervention group vs control group) was also included to evaluate the effect of each treatment. Residual depressive symptoms at discharge and its interaction with the study group were also added. To detect remission rates, we calculated the Reliable Change Index (RCI; based on the approach of Jacobson and Truax [42]). The RCI was calculated for the change of the primary outcome (BDI-II) between baseline (T0) and the respective follow-ups (T1, T2, or T3) for both groups (intervention group and control group) as follows: (1) Remission was defined as a BDI-II reduction of at least 8 and a total score below 14 (=normal range). (2) Improved but not recovered was defined as a BDI-II score reduction of at least 8 points but a total score above 14. (3) No reliable change was defined as a BDI-II score neither increased above, nor reduced by more than 8 points. (4) Deteriorated was defined as a BDI-II increase above 8 and a total score above 14. Data analyses were performed with SPSS version 23 (IBM SPSS Statistics 23 [44]).

Power Analysis

As described in the study protocol [30], the study had a power of 0.97 to detect an effect size of d=0.50 or higher with a sample size of N=230.

Randomization

As delineated in the study by Zwerenz et al [28], randomization (intervention group and control group) was performed using the software Research Randomizer [45] by the Study Center of Mental Disorders of the University Medical Center Mainz (block randomization at a ratio of 1:1) faxing the assignment to the study assistant at the clinic, who performed the assignment, accordingly.

Results

Primary and Secondary Outcomes

Table 1 reports means and SDs of the observed scores at baseline and the estimated scores at follow-up (T3) for the primary and secondary outcomes.

At follow-up (T3), intervention group and control group significantly improved regarding depressive symptoms assessed with the BDI-II (Table 1). Within-group effect size was high in the intervention group and moderate in the control group. The between-group comparison showed a medium effect size (P<.001; d=0.58). As can also be seen in Figure 1, the intervention group shows at any point in time after T0 a stronger reduction in the BDI-II than the control group.

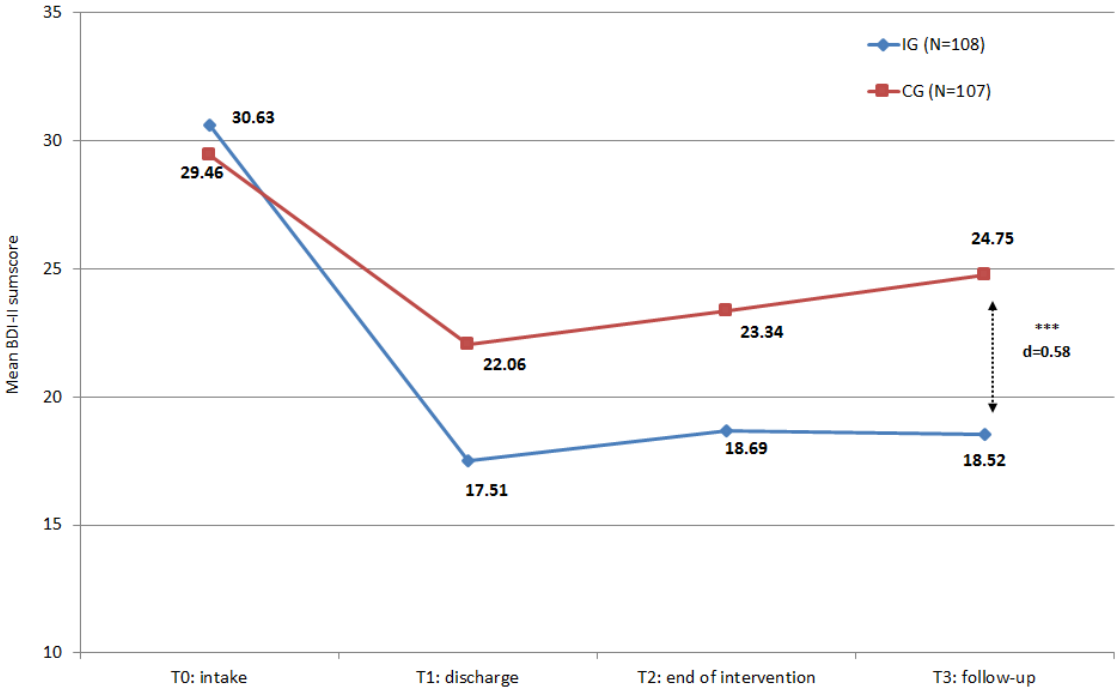
Table 2 reports test statistics as well as effect sizes for the primary and secondary outcomes at follow-up (T3) after 6 months compared with study intake (T0).

Table 1. Descriptive statistics of primary and secondary outcomes at 6-month follow-up (T3) of intervention (N=108) and control group (N=107), compared with study admission (T0).

Outcome criteria	Baseline (T0[a]), mean (SD)	Follow-Up (T3[b]), mean[c] (SD)
BDI-II[d]		
IG[e]	30.63 (9.39)	18.52 (10.78)
CG[f]	29.46 (9.50)	24.75 (10.74)
PHQ-9[g]		
IG	14.86 (4.88)	10.37 (5.54)
CG	14.30 (5.23)	13.13 (5.70)
GAD-7[h]		
IG	11.59 (4.28)	8.24 (4.68)
CG	11.57 (5.08)	10.38 (4.68)
EUROHIS[i]		
IG	1.62 (0.57)	2.13 (0.68)
CG	1.68 (0.59)	1.84 (0.66)
RSE[j]		
IG	15.11 (6.68)	17.78 (6.60)
CG	15.40 (6.80)	15.75 (6.60)
DAS[k]		
IG	149.96 (37.05)	139.78 (34.26)
CG	158.42 (38.26)	145.82 (34.78)
WAI[l]		
IG	25.41 (2.87)	26.47 (3.73)
CG	25.28 (2.60)	26.30 (4.55)

[a]T0: allocation to intervention (baseline).
[b]T3: follow-up 6 months after baseline.
[c]Estimated means.
[d]BDI-II: Beck Depression Inventory.
[e]IG: intervention group.
[f]CG: control group.
[g]PHQ-9: Patient Health Questionnaire-9.
[h]GAD-7: Generalized Anxiety Disorder-7.
[i]EUROHIS: European Health Interview Survey.
[j]RSE: Rosenberg Self-Esteem Scale.
[k]DAS: Dysfunctional Attitude Scale.
[l]WAI: Work Ability Index.

Figure 1. Primary outcome (BDI-II) in the course of time. BDI-II: Beck Depression Inventory II; CG: control group; IG: intervention group; T0: allocation to intervention (baseline); T3: follow-up 6 months after baseline.



There were no significant differences at baseline (T0) between intervention group and control group for any of the secondary outcome measures. A lower symptom load at follow-up (T3) in the intervention group in comparison with the control group could also be observed for the secondary outcomes. Analyses revealed statistically significant between group differences and low-to-moderate effect sizes for depressive symptoms assessed with the PHQ-9 (P<.001; d=0.49), anxiety (GAD-7; P<.001; d=0.46); quality of life (EUROHIS-QOL 8; P=.002; d=0.43), and self-esteem measured by the RSE scale (P=.02; d=0.31). There were no significant differences comparing both groups concerning dysfunctional attitudes assessed with the DAS (P=.34; d=0.18) and work ability (WAI; P=.45; d=0.04). Furthermore, within-group comparisons showed that both groups benefitted not only regarding depressive symptoms measured by the PHQ-9 with an effect size of d=0.80 in the intervention group and d=0.22 in the control group but also for anxiety (intervention group: d=0.69, control group: d=0.24) and quality of life (intervention group: d=0.71, control group: d=0.25).

Table 2. Between-group and within-group comparisons for primary and secondary outcomes.

Outcome criteria	T3[a]			T0[b]-T3		
	Between-group comparisons[c]			Within-group comparisons[d]		
	F (*df*)	*P* value[e]	*d*	*t* (*df*); *r*	*P* value	*d*
BDI-II[f]	18.87 (1,212)	<.001	0.58	—[g]	—	—
IG[h]	—	—	—	10.21 (1665); 0.42	<.001	1.06
CG[i]	—	—	—	4.94 (1671); 0.57	<.001	0.44
PHQ-9[j]	13.733 (1,212)	<.001	0.49	—	—	—
IG	—	—	—	7.488 (13863); 0.38	<.001	0.80
CG	—	—	—	2.136 (1088); 0.42	<.001	0.23
GAD-7[k]	12.802 (1,212)	<.001	0.46	—	—	—
IG	—	—	—	6.497 (6851); 0.39	<.001	0.69
CG	—	—	—	2.252 (19067); 0.38	<.001	0.24
EUROHIS[l]	9.557 (1,212)	<.001	0.43	—	—	—
IG	—	—	—	6.895 (1095); 0.43	<.001	0.71
CG	—	—	—	2.488 (866); 0.48	<.001	0.25
RSE[m]	5.427 (1,212)	.02	0.31	—	—	—
IG	—	—	—	3.585 (7165); 0.44	<.001	0.37
CG	—	—	—	0.589 (4437), 0.45	.55	0.06
DAS[n]	0.955 (1,212)	.34	0.18	—	—	—
IG	—	—	—	3.917 (34211); 0.62	<.001	0.33
CG	—	—	—	2.558 (7731); 0.52	.01	0.24
WAI[o]	0.676 (1,212)	.45	0.04	—	—	—
IG	—	—	—	2.558 (99); 0.22	.01	0.31
CG	—	—	—	1.948 (31); 0.04	.06	0.26

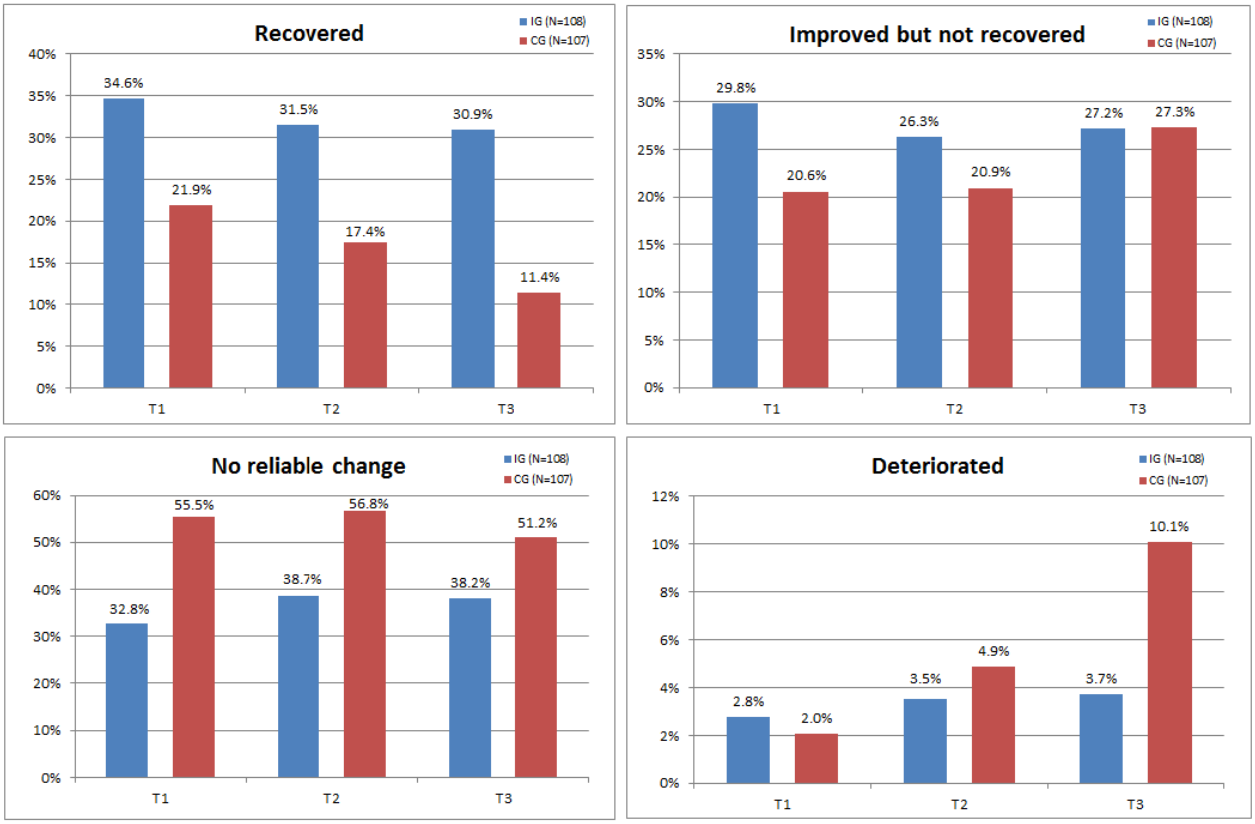
[a]T3: Follow-up 6 months after baseline.
[b]T0: Allocation to intervention (baseline).
[c]Analyses of covariance with baseline as covariate.
[d]Paired samples *t* tests.
[e]Level of significance .05.
[f]BDI-II: Beck Depression Inventory II.
[g]Not applicable.
[h]IG: intervention group.
[i]CG: control group.
[j]PHQ-9: Patient Health Questionnaire-9.
[k]GAD-7: Generalized Anxiety Disorde-7.
[l]EUROHIS: European Health Interview Survey.
[m]RSE: Rosenberg Self-Esteem Scale.
[n]DAS: Dysfunctional Attitude Scale.
[o]WAI: Work Ability Index.

Remission, Improvement, and Deterioration

Comparisons between intervention group and control group are shown in Figure 2. As can be seen in Figure 2, the rate of participants experiencing a remission or an improvement was significantly higher in the intervention group than in the control group at each point of measurement (T1: χ^2_3=11.5; *P*=.01; *d*=0.36; T2: χ^2_3=9.3; *P*=.03; *d*=0.31; T3: χ^2_3=15.2; *P*=.002; *d*=0.55). The gap even widened over time with an NNT of 7.84

at discharge (T1), 7.09 at the end of intervention (T2), and 5.12 at follow-up (T3).

Figure 2. Recovery, improvement, and deterioration in the course of time. BDI-II: Beck Depression Inventory II; CG: control group; IG: intervention group; T1: discharge; T2: end of intervention; T3: follow-up 6 months after baseline.



Predictors of Depressive Symptoms at Follow-Up 6 Months Later

Following the study protocol, multivariate analyses were performed to test the effects of postulated predictors on depressive symptoms at follow-up 6 months later (T3). These were depressive symptoms at study intake and at discharge, utilization of other treatments at follow-up, childhood trauma, and structural characteristics. GLM (R^2=0.688; $F_{118,200}$=4.813; P<.04, η_p^2=0.040) based on imputed data revealed that depressive symptoms at discharge from hospital treatment was the strongest predictor (Table 3). Interestingly, baseline depressive symptoms played no additional role. Group membership (Deprexis vs control group) was an additional predictor of depressive symptoms at follow-up, and there was a trend to an interaction between depressive symptoms at discharge and group. Other variables which had been posited in our study protocol as potential predictors (psychotherapeutic treatment at follow-up, childhood trauma, structural characteristics) as well as sex and age played no role as predictors.

Table 3. Predictors of depressive symptoms at 6-month follow-up (T3) based on generalized linear model (N=200).

Predictors	*F* (*df*)	*P* value[a]	η_p^2 [b]
Age	0.460 (1)	.25	0.004
Sex	0.316 (1)	.28	0.003
CTQ[c] at T0[d]	0.991 (1)	.16	0.008
OPD-SFK[e] at T0	0.454 (1)	.25	0.004
Utilization of other treatments at follow-up	0.294 (1)	.30	0.002
BDI-II[f] at T0	1.285 (1)	.26	0.011
BDI-II at T1[g]	2.630 (46)	<.001	0.506
Group	3.243 (1)	.04	0.027
BDI-II at T1 X Group	1.267 (28)	.10	0.231

[a]Level of significance .05, 1-tailed test.

[b]Partial eta square.

[c]CTQ: Childhood Trauma Questionnaire.

[d]T0: Study admission (baseline).

[e]OPD-SFK: Operationalized Psychodynamic Diagnosis.

[f]BDI-II: Beck Depression Inventory.

[g]T1: Discharge of inpatient treatment.

Discussion

Principal Findings

Following our first publication to the efficacy of Deprexis offered as an add-on to inpatient psychotherapy [28], this study investigated effects on the course of depressive symptoms of the Web-based self-management program 6 months after study inclusion. Participation in the adjunct Web-based self-management program improved the course of depressive symptoms during the follow-up period. The gap of effectiveness between the intervention group and the control group, which received access to Web-based information about depression and depressive symptoms in addition to inpatient psychotherapy, even widened over the course of the study. This was reflected in an increase of the between group effect size from *d*=0.44 to *d*=0.58. When we differentiated between different categories of outcome, the proportion of remission of 35% in the Deprexis group at discharge was almost maintained at follow-up (31%), whereas remission in the control group declined from 22% to 11% in the same period. Comparable proportions of patients achieved substantial improvements but did not fulfill criteria of remission (about 27%) at follow-up. More than twice as many participants (10% vs 4%) deteriorated in the control group compared with the intervention group 6 months after study inclusion. Correspondingly, the NNT improved in favor of the intervention group (from 7.84 at discharge from inpatient treatment to 5.12 at follow-up). This result, that 1 in every 5 to 8 patients benefitted from the program, is not only comparable with other studies on self-guided interventions [46] but also with results from face-to-face therapies [47].

Additional and moderate improvements were found in Deprexis compared with the control group regarding the PHQ-9 measure of depressive symptoms, generalized anxiety (GAD-7), and quality of life (EUROHIS-QOL), all of which also improved in the control group. Self-esteem (RSE) only improved moderately in the intervention group but not the control group. Dysfunctional attitudes (DAS) improved in both groups, whereas WAI was significant in the intervention (trend in the control group), and there were no differential benefits.

Consistent with previous studies [26], residual depressive symptoms at termination of inpatient treatment were the strongest predictor of depressive symptoms at follow-up, in addition to a small group effect in favor of the intervention. Baseline scores of depressive symptoms, childhood trauma, psychic structure, and demographic characteristics played no additional role. Thus, superiority of the intervention versus the control condition rests mainly on the fact that the intervention reduces residual depressive symptoms compared with controls. As indicated by the trend to an interaction, continued use of the program may further reduce depressive symptoms during follow-up. However, this is a very small effect at best.

Inpatient treatment usually applies to complex cases of depression with different mental and somatic comorbidities [48]. Resonating with other findings [26], inpatient psychotherapy should strive not only to improve depressive symptoms scores (among other goals) but also to reduce residual depressive symptoms further. The relationship of symptom reduction and duration, respectively, dose of treatment is considered complex because of different patient trajectories of change and therapist influences [49]. Compared with 61.8 days in a multicenter study conducted by Franz et al [23] from 2007 to 2011 and 10 weeks in the previous trial of Zeek et al [25], the total duration of 40 days of inpatient treatment was relatively short. Cutting down treatment duration in the medical system to save cost may lead to higher rates of relapse or recurrence.

Strengths and Limitations

The high effects of our adjunct to inpatient multimodal psychotherapy are consistent with a recent review, in which evidence was shown that Web-based interventions are efficacious for maintaining treatment gains and prevent relapse [50]. This leads to the question, what barriers have to be overcome and what are the challenges for integrating new technologies and interventions into an existing treatment setting, such as, for example, low acceptance of Web-based interventions in the treatment team [51,52], technological barriers, or low acceptance and compliance within patients [53].

In view of chronic trajectories of depressive symptoms, the short follow-up interval of about 3 months (mean 96.59 days, SD 20.21 days) following termination of the program must be considered a drawback. However, our data underline that even the brief time period of the first months after discharge sets the stage for deterioration for many patients—notwithstanding continued outpatient care. Although the study raises important economic issues, health economic data were not collected. Another limitation is that we mainly relied on self-reports and have no clinical diagnostics after discharge from inpatient treatment. Furthermore, we do not know exactly what kind of treatment and what treatment intensity patients received after discharge from the hospital and of course this could have an influence on the course of depressive symptoms. But after all, we could control for the 1 item that asked if any kind of treatment followed inpatient treatment (yes/no), which was not a significant predictor of depressive symptoms at follow-up.

Although the Web-based self-help program Deprexis has been demonstrated to be efficacious [11], the challenge is how to integrate it into regular mental health care. To our knowledge, we were the first to compare Deprexis as an add-on to inpatient psychotherapy to an active control group receiving Web-based psychoeducational information in an RCT. Compliance and retention were very good with dropout rates between 3% and 4% between the different timepoints of the study and no differences in dropout rates between intervention group and control group [28]. We, therefore, presume that a structured inpatient treatment format is suitable for adding adjunct Web-based self-help. However, as blended therapy is currently regarded as promising to improve psychotherapeutic treatment, the question arises how implementation could take place and what challenges have to be overcome. In a qualitative study of therapist's perspective on blended psychotherapy for depression [54], the most frequent barriers reported were limited customizability and autonomy of decisions concerning blended treatment, disease-related contraindications, negative affect because of technical problems, the limitation of face-to-face sessions as a consequence of blending the therapy, or the impairment of therapeutic alliance because of technical problems. Facilitators were also mentioned but less frequently, such as, motivation and willingness of patients for innovative interventions, the possibility for patients to use Web-based contents between face-to-face sessions, or the contemporary treatment possibility which could also close the treatment gap [54].

Conclusions

Tapping into self-help resources of patients by adding Web-based self-help is a promising way to improve long-term outcomes. There is a need for more studies, but if other studies also conclude that Web-based interventions as an add-on or blended treatment are effective, the challenge will be to implement them in the health care system so that reimbursement is possible. An ongoing development, regular updating, and securing of the latest technical standards are associated with corresponding costs. Therefore, it must be investigated who can bear these costs and how much they are. However, if that succeeds, efficacious Web-based interventions could be implemented on a broader basis to improve the benefits and sustainability of face-to-face treatments.

Acknowledgments

The authors are grateful to all the participating patients and the staff in the cooperating clinic for their support in patient recruitment without whom this study would not be possible (in alphabetical order), especially Johanna Kraus, Dipl-Sozpäd Annerose Vorndran, Dipl-Sozpäd Hiltrud Zajac, as well as our student research assistants Svenja Bidmon (BSc Psych), Tanja Darwiesh (MSc Psych), and Hanna Merz (BSc Psych). The study was funded by Rhön-Klinikum AG Bad Neustadt/Saale.

Authors' Contributions

MEB and RZ did the first draft of the manuscript. RZ, CB, AT, JB, and MS did the final draft of the manuscript and revised it critically for its intellectual content. RZ, JB, MS, RJK, and MEB substantially contributed to the conception and the design of the study. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

References

1. Wittchen HU, Jacobi F, Rehm J, Gustavsson A, Svensson M, Jönsson B, et al. The size and burden of mental disorders and other disorders of the brain in Europe 2010. Eur Neuropsychopharmacol 2011 Sep;21(9):655-679. [doi: 10.1016/j.euroneuro.2011.07.018] [Medline: 21896369]

2. Wang J, Wu X, Lai W, Long E, Zhang X, Li W, et al. Prevalence of depression and depressive symptoms among outpatients: a systematic review and meta-analysis. BMJ Open 2017 Aug 23;7(8):e017173 [FREE Full text] [doi: 10.1136/bmjopen-2017-017173] [Medline: 28838903]
3. Patel V, Maj M, Flisher AJ, de Silva MJ, Koschorke M, Prince M, WPA Zonal and Member Society Representatives. Reducing the treatment gap for mental disorders: a WPA survey. World Psychiatry 2010 Oct;9(3):169-176 [FREE Full text] [doi: 10.1002/j.2051-5545.2010.tb00305.x] [Medline: 20975864]
4. Rogers MA, Lemmen K, Kramer R, Mann J, Chopra V. Internet-delivered health interventions that work: systematic review of meta-analyses and evaluation of website availability. J Med Internet Res 2017 Mar 24;19(3):e90 [FREE Full text] [doi: 10.2196/jmir.7111] [Medline: 28341617]
5. Cuijpers P, Kleiboer A, Karyotaki E, Riper H. Internet and mobile interventions for depression: opportunities and challenges. Depress Anxiety 2017 Jul;34(7):596-602. [doi: 10.1002/da.22641] [Medline: 28471479]
6. Kok G, Burger H, Riper H, Cuijpers P, Dekker J, van Marwijk H, et al. The three-month effect of mobile internet-based cognitive therapy on the course of depressive symptoms in remitted recurrently depressed patients: results of a randomized controlled trial. Psychother Psychosom 2015;84(2):90-99. [doi: 10.1159/000369469] [Medline: 25721915]
7. Holländare F, Johnsson S, Randestad M, Tillfors M, Carlbring P, Andersson G, et al. Randomized trial of internet-based relapse prevention for partially remitted depression. Acta Psychiatr Scand 2011 Oct;124(4):285-294. [doi: 10.1111/j.1600-0447.2011.01698.x] [Medline: 21401534]
8. Twomey C, O'Reilly G. Effectiveness of a freely available computerised cognitive behavioural therapy programme (MoodGYM) for depression: Meta-analysis. Aust N Z J Psychiatry 2017 Mar;51(3):260-269. [doi: 10.1177/0004867416656258] [Medline: 27384752]
9. Meyer B, Berger T, Caspar F, Beevers CG, Andersson G, Weiss M. Effectiveness of a novel integrative online treatment for depression (Deprexis): randomized controlled trial. J Med Internet Res 2009 May 11;11(2):e15 [FREE Full text] [doi: 10.2196/jmir.1151] [Medline: 19632969]
10. Karyotaki E, Riper H, Twisk J, Hoogendoorn A, Kleiboer A, Mira A, et al. Efficacy of self-guided internet-based cognitive behavioral therapy in the treatment of depressive symptoms: a meta-analysis of individual participant data. JAMA Psychiatry 2017 Apr 1;74(4):351-359. [doi: 10.1001/jamapsychiatry.2017.0044] [Medline: 28241179]
11. Twomey C, O'Reilly G, Meyer B. Effectiveness of an individually-tailored computerised CBT programme (Deprexis) for depression: a meta-analysis. Psychiatry Res 2017 Oct;256:371-377. [doi: 10.1016/j.psychres.2017.06.081] [Medline: 28686935]
12. Klein JP, Berger T, Schröder J, Späth C, Meyer B, Caspar F, et al. Effects of a psychological internet intervention in the treatment of mild to moderate depressive symptoms: results of the EVIDENT study, a randomized controlled trial. Psychother Psychosom 2016;85(4):218-228. [doi: 10.1159/000445355] [Medline: 27230863]
13. Klein JP, Späth C, Schröder J, Meyer B, Greiner W, Hautzinger M, et al. Time to remission from mild to moderate depressive symptoms: one year results from the EVIDENT-study, an RCT of an internet intervention for depression. Behav Res Ther 2017 Oct;97:154-162. [doi: 10.1016/j.brat.2017.07.013] [Medline: 28797829]
14. Gilbody S, Littlewood E, Hewitt C, Brierley G, Tharmanathan P, Araya R, REEACT Team. Computerised cognitive behaviour therapy (cCBT) as treatment for depression in primary care (REEACT trial): large scale pragmatic randomised controlled trial. Br Med J 2015 Nov 11;351:h5627 [FREE Full text] [doi: 10.1136/bmj.h5627] [Medline: 26559241]
15. Klein NS, Bockting CL, Wijnen B, Kok GD, van Valen E, Riper H, et al. Economic evaluation of an internet-based preventive cognitive therapy with minimal therapist support for recurrent depression: randomized controlled trial. J Med Internet Res 2018 Nov 26;20(11):e10437 [FREE Full text] [doi: 10.2196/10437] [Medline: 30478021]
16. Kleiboer A, Smit J, Bosmans J, Ruwaard J, Andersson G, Topooco N, et al. European COMPARative effectiveness research on blended depression treatment versus treatment-as-usual (E-COMPARED): study protocol for a randomized controlled, non-inferiority trial in eight European countries. Trials 2016 Aug 3;17(1):387 [FREE Full text] [doi: 10.1186/s13063-016-1511-1] [Medline: 27488181]
17. Nakao S, Nakagawa A, Oguchi Y, Mitsuda D, Kato N, Nakagawa Y, et al. Web-based cognitive behavioral therapy blended with face-to-face sessions for major depression: randomized controlled trial. J Med Internet Res 2018 Sep 21;20(9):e10743 [FREE Full text] [doi: 10.2196/10743] [Medline: 30249583]
18. Schuster R, Sigl S, Berger T, Laireiter AR. Patients' experiences of web- and mobile-assisted group therapy for depression and implications of the group setting: qualitative follow-up study. JMIR Ment Health 2018 Jul 11;5(3):e49 [FREE Full text] [doi: 10.2196/mental.9613] [Medline: 29997106]
19. Schuster R, Leitner I, Carlbring P, Laireiter AR. Exploring blended group interventions for depression: randomised controlled feasibility study of a blended computer- and multimedia-supported psychoeducational group intervention for adults with depressive symptoms. Internet Interv 2017 Jun;8:63-71 [FREE Full text] [doi: 10.1016/j.invent.2017.04.001] [Medline: 30135830]
20. McCrone P, Knapp M, Proudfoot J, Ryden C, Cavanagh K, Shapiro DA, et al. Cost-effectiveness of computerised cognitive-behavioural therapy for anxiety and depression in primary care: randomised controlled trial. Br J Psychiatry 2004 Jul;185:55-62. [doi: 10.1192/bjp.185.1.55] [Medline: 15231556]

21. Kenter RM, van de Ven PM, Cuijpers P, Koole G, Niamat S, Gerrits RS, et al. Costs and effects of internet cognitive behavioral treatment blended with face-to-face treatment: results from a naturalistic study. Internet Interv 2015 Mar;2(1):77-83. [doi: 10.1016/j.invent.2015.01.001]
22. NVL - Nationale VersorgungsLeitlinien — ÄZQ. 2015. Unipolare Depression: S3-Leitlinie und Nationale Versorgungsleitlinie (Nvl) Unipolare Depression, 2. Auflage URL: https://www.leitlinien.de/nvl/depression [accessed 2019-08-07]
23. Franz M, Seidler D, Jenett D, Schäfer R, Wutzler U, Kämmerer W, et al. [Psychodynamically oriented inpatient psychotherapy of depressive disorders - first findings from a naturalistic, multicenter study]. Z Psychosom Med Psychother 2015;61(1):19-35. [doi: 10.13109/zptm.2015.61.1.19] [Medline: 25831982]
24. Liebherz S, Rabung S. Do patients' symptoms and interpersonal problems improve in psychotherapeutic hospital treatment in Germany? A systematic review and meta-analysis. PLoS One 2014;9(8):e105329 [FREE Full text] [doi: 10.1371/journal.pone.0105329] [Medline: 25141289]
25. Zeeck A, von Wietersheim J, Weiß H, Scheidt CE, Völker A, Helesic A, et al. Symptom course in inpatient and day clinic treatment of depression: results from the INDDEP-study. J Affect Disord 2015 Nov 15;187:35-44. [doi: 10.1016/j.jad.2015.07.025] [Medline: 26318269]
26. Kuehner C, Huffziger S. Factors predicting the long-term illness course in a cohort of depressed inpatients. Eur Arch Psychiatry Clin Neurosci 2013 Aug;263(5):413-423. [doi: 10.1007/s00406-012-0379-7] [Medline: 23108435]
27. Judd LL, Akiskal HS, Maser JD, Zeller PJ, Endicott J, Coryell W, et al. Major depressive disorder: a prospective study of residual subthreshold depressive symptoms as predictor of rapid relapse. J Affect Disord 1998 Sep;50(2-3):97-108. [doi: 10.1016/S0165-0327(98)00138-4] [Medline: 9858069]
28. Zwerenz R, Becker J, Knickenberg RJ, Siepmann M, Hagen K, Beutel ME. Online self-help as an add-on to inpatient psychotherapy: efficacy of a new blended treatment approach. Psychother Psychosom 2017;86(6):341-350. [doi: 10.1159/000481177] [Medline: 29131090]
29. Vittengl JR, Clark LA, Dunn TW, Jarrett RB. Reducing relapse and recurrence in unipolar depression: a comparative meta-analysis of cognitive-behavioral therapy's effects. J Consult Clin Psychol 2007 Jun;75(3):475-488 [FREE Full text] [doi: 10.1037/0022-006X.75.3.475] [Medline: 17563164]
30. Zwerenz R, Becker J, Knickenberg RJ, Hagen K, Dreier M, Wölfling K, et al. Enhancing inpatient psychotherapeutic treatment with online self-help: study protocol for a randomized controlled trial. Trials 2015 Mar 17;16:98 [FREE Full text] [doi: 10.1186/s13063-015-0620-6] [Medline: 25872505]
31. SoSci Survey. 2014. URL: https://www.soscisurvey.de/ [accessed 2018-08-07]
32. Wang YP, Gorenstein C. Psychometric properties of the Beck Depression Inventory-II: a comprehensive review. Braz J Psychiatry 2013;35(4):416-431 [FREE Full text] [doi: 10.1590/1516-4446-2012-1048] [Medline: 24402217]
33. Kühner C, Bürger C, Keller F, Hautzinger M. [Reliability and validity of the revised Beck Depression Inventory (BDI-II). Results from German samples]. Nervenarzt 2007 Jun;78(6):651-656. [doi: 10.1007/s00115-006-2098-7] [Medline: 16832698]
34. Gelaye B, Williams MA, Lemma S, Deyessa N, Bahretibeb Y, Shibre T, et al. Validity of the patient health questionnaire-9 for depression screening and diagnosis in east Africa. Psychiatry Res 2013 Dec 15;210(2):653-661 [FREE Full text] [doi: 10.1016/j.psychres.2013.07.015] [Medline: 23972787]
35. Spitzer RL, Kroenke K, Williams JB, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med 2006 May 22;166(10):1092-1097. [doi: 10.1001/archinte.166.10.1092] [Medline: 16717171]
36. da Rocha NS, Power MJ, Bushnell DM, Fleck MP. The EUROHIS-QOL 8-item index: comparative psychometric properties to its parent WHOQOL-BREF. Value Health 2012 May;15(3):449-457 [FREE Full text] [doi: 10.1016/j.jval.2011.11.035] [Medline: 22583455]
37. von Collani G, Herzberg PY. Eine revidierte fassung der deutschsprachigen skala zum selbstwertgefühl von rosenberg. Zeitschrift für Differentielle und Diagnostische Psychologie 2003 Mar;24(1):3-7. [doi: 10.1024//0170-1789.24.1.3]
38. de Graaf LE, Roelofs J, Huibers MJ. Measuring dysfunctional attitudes in the general population: the dysfunctional attitude scale (Form A) revised. Cognit Ther Res 2009 Aug;33(4):345-355 [FREE Full text] [doi: 10.1007/s10608-009-9229-y] [Medline: 19623267]
39. Ilmarinen J. The work ability index (WAI). Occup Med 2006 Oct 17;57(2):160. [doi: 10.1093/occmed/kqm008]
40. Wingenfeld K, Spitzer C, Mensebach C, Grabe HJ, Hill A, Gast U, et al. [The German version of the childhood trauma questionnaire (CTQ): preliminary psychometric properties]. Psychother Psychosom Med Psychol 2010 Nov;60(11):442-450. [doi: 10.1055/s-0030-1247564] [Medline: 20200804]
41. Ehrenthal JC, Dinger U, Horsch L, Komo-Lang M, Klinkerfuss M, Grande T, et al. [The OPD structure questionnaire (OPD-SQ): first results on reliability and validity]. Psychother Psychosom Med Psychol 2012 Jan;62(1):25-32. [doi: 10.1055/s-0031-1295481] [Medline: 22271173]
42. Jacobson NS, Truax P. Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. J Consult Clin Psychol 1991 Feb;59(1):12-19. [doi: 10.1016/S0005-7894(81)80119-0] [Medline: 2002127]
43. Cook RJ, Sackett DL. The number needed to treat: a clinically useful measure of treatment effect. Br Med J 1995 Feb 18;310(6977):452-454 [FREE Full text] [doi: 10.1136/bmj.310.6977.452] [Medline: 7873954]
44. IBM Knowledge Center. 2014. Downloading IBM SPSS Statistics 23 URL: https://www-01.ibm.com/support/docview.wss?uid=swg24038592 [accessed 2018-08-07]

45. Research Randomizer. 2013. URL: http://www.randomizer.org/ [accessed 2018-08-07]
46. Cuijpers P, Donker T, Johansson R, Mohr DC, van Straten A, Andersson G. Self-guided psychological treatment for depressive symptoms: a meta-analysis. PLoS One 2011;6(6):e21274 [FREE Full text] [doi: 10.1371/journal.pone.0021274] [Medline: 21712998]
47. Cuijpers P, van Straten A, Bohlmeijer E, Hollon S, Andersson G. The effects of psychotherapy for adult depression are overestimated: a meta-analysis of study quality and effect size. Psychol Med 2010 Feb;40(2):211-223. [doi: 10.1017/S0033291709006114] [Medline: 19490745]
48. Zwerenz R, Becker J, Johansson R, Frederick RJ, Andersson G, Beutel ME. Transdiagnostic, psychodynamic web-based self-help intervention following inpatient psychotherapy: results of a feasibility study and randomized controlled trial. JMIR Ment Health 2017 Oct 16;4(4):e41 [FREE Full text] [doi: 10.2196/mental.7889] [Medline: 29038094]
49. Stulz N, Lutz W, Kopta SM, Minami T, Saunders SM. Dose-effect relationship in routine outpatient psychotherapy: does treatment duration matter? J Couns Psychol 2013 Oct;60(4):593-600. [doi: 10.1037/a0033589] [Medline: 23815633]
50. Hennemann S, Farnsteiner S, Sander L. Internet- and mobile-based aftercare and relapse prevention in mental disorders: a systematic review and recommendations for future research. Internet Interv 2018 Dec;14:1-17 [FREE Full text] [doi: 10.1016/j.invent.2018.09.001] [Medline: 30510909]
51. Hennemann S, Beutel ME, Zwerenz R. ['Morbus Google' vs e-health: a qualitative study of acceptance and implementation of online-aftercare in inpatient rehabilitation]. Rehabilitation (Stuttg) 2018 Feb;57(1):14-23. [doi: 10.1055/s-0043-102150] [Medline: 28591900]
52. Hennemann S, Witthöft M, Bethge M, Spanier K, Beutel ME, Zwerenz R. Acceptance and barriers to access of occupational e-mental health: cross-sectional findings from a health-risk population of employees. Int Arch Occup Environ Health 2018 Apr;91(3):305-316. [doi: 10.1007/s00420-017-1280-5] [Medline: 29189895]
53. Hennemann S, Beutel ME, Zwerenz R. Drivers and barriers to acceptance of web-based aftercare of patients in inpatient routine care: a cross-sectional survey. J Med Internet Res 2016 Dec 23;18(12):e337 [FREE Full text] [doi: 10.2196/jmir.6003] [Medline: 28011445]
54. Titzler I, Saruhanjan K, Berking M, Riper H, Ebert DD. Barriers and facilitators for the implementation of blended psychotherapy for depression: a qualitative pilot study of therapists' perspective. Internet Interv 2018 Jun;12:150-164 [FREE Full text] [doi: 10.1016/j.invent.2018.01.002] [Medline: 30135779]

Abbreviations

BDI-II: Beck Depression Inventory-II
DAS: Dysfunctional Attitude Scale
EUROHIS-QOL 8: European Health Interview Survey Quality of Life 8-Item Index
GAD-7: generalized anxiety disorder-7
GLM: generalized linear model
NNT: number needed to treat
PHQ-9: Patient Health Questionnaire-9
RCI: Reliable Change Index
RCT: randomized controlled trial
RSE: Rosenberg Self-Esteem Scale
TAU: Treatment as Usual
WAI: Work Ability Index

Edited by G Eysenbach; submitted 14.02.19; peer-reviewed by B Meyer, C Chen, T Krieger; comments to author 28.03.19; revised version received 20.06.19; accepted 07.07.19; published 24.10.19.

Please cite as:
Zwerenz R, Baumgarten C, Becker J, Tibubos A, Siepmann M, Knickenberg RJ, Beutel ME
Improving the Course of Depressive Symptoms After Inpatient Psychotherapy Using Adjunct Web-Based Self-Help: Follow-Up Results of a Randomized Controlled Trial
J Med Internet Res 2019;21(10):e13655
URL: https://www.jmir.org/2019/10/e13655
doi:10.2196/13655
PMID:31651403

Original Paper

Utilization, Safety, and Technical Performance of a Telemedicine System for Prehospital Emergency Care: Observational Study

Marc Felzen[1], MD, Dr med; Stefan Kurt Beckers[1], MD, Dr med; Felix Kork[1], MD, Dr med, MSc; Frederik Hirsch[1], MD, Dr med; Sebastian Bergrath[2], MD, Dr med; Anja Sommer[3], MSc; Jörg Christian Brokmann[4], MD, Dr med; Michael Czaplik[1], MD, PhD; Rolf Rossaint[1], MD, Prof Dr

[1]Department of Anesthesiology, Medical Faculty, RWTH Aachen University, Aachen, Germany

[2]Emergency Department, Maria Hilf-Hospital Moenchengladbach, Moenchengladbach, Germany

[3]Department of Health, Ethics and Society, Care and Public Health Research Institute, Maastricht University, Maastricht, Netherlands

[4]Emergency Department, Medical Faculty, RWTH Aachen University, Aachen, Germany

Corresponding Author:
Marc Felzen, MD, Dr med
Department of Anesthesiology
Medical Faculty
RWTH Aachen University
Pauwelsstr 30
Aachen, 52074
Germany
Phone: 49 2418088179
Email: mfelzen@ukaachen.de

Abstract

Background: As a consequence of increasing emergency medical service (EMS) missions requiring an EMS physician on site, we had implemented a unique prehospital telemedical emergency service as a new structural component to the conventional physician-based EMS in Germany.

Objective: We sought to assess the utilization, safety, and technical performance of this telemedical emergency service.

Methods: We conducted a retrospective analysis of all primary emergency missions with telemedical consultation of an EMS physician in the City of Aachen (250,000 inhabitants) during the first 3 operational years of our tele-EMS system. Main outcome measures were the number of teleconsultations, number of complications, and number of transmission malfunctions during teleconsultations.

Results: The data of 6265 patients were analyzed. The number of teleconsultations increased during the run-in period of four quarters toward full routine operation from 152 to 420 missions per quarter. When fully operational, around the clock, and providing teleconsultations to 11 mobile ambulances, the number of teleconsultations further increased by 25.9 per quarter (95% CI 9.1-42.6; *P*=.009). Only 6 of 6265 patients (0.10%; 95% CI 0.04%-0.21%) experienced adverse events, all of them not inherent in the system of teleconsultations. Technical malfunctions of single transmission components occurred from as low as 0.3% (95% CI 0.2%-0.5%) during two-way voice communications to as high as 1.9% (95% CI 1.6%-2.3%) during real-time vital data transmissions. Complete system failures occurred in only 0.3% (95% CI 0.2%-0.6%) of all teleconsultations.

Conclusions: The Aachen prehospital EMS is a frequently used, safe, and technically reliable system to provide medical care for emergency patients without an EMS physician physically present. Noninferiority of the tele-EMS physician compared with an on-site EMS physician needs to be demonstrated in a randomized trial.

(J Med Internet Res 2019;21(10):e14907) doi:10.2196/14907

KEYWORDS

emergency medicine; ambulances; telemedicine; quality of care; eHealth

Introduction

Background

German emergency medical services (EMSs) comprise a system of cooperating on-site paramedics and physicians who are separately dispatched to the patient depending on the severity of the emergency. Over recent years, EMSs have faced a serious problem: the number of missions requiring on-site EMS physicians has continuously been rising [1], making dispatchable EMS physicians a scarce resource.

Several factors have facilitated this development. First, by German law, only physicians are entitled to prescribe drugs. Paramedics are allowed to administer drugs only if (1) the situation is life-threatening, (2) the administration and dosing are predefined in a standard operating procedure (SOP), (3) a less invasive measure to achieve the same effect is not available, and (4) an EMS physician has already been dispatched to the scene. Hence, EMS physicians are regularly dispatched to patients with non–life-threatening conditions (eg, being in pain) and are—during that time—indispensable for other potentially more severe emergency missions. Second, in one-fifth of emergency missions involving an EMS physician, the EMS physician is subsequently requested by the paramedics on site and was not dispatched initially [2]. As a consequence, the treatment of these patients is distinctly delayed. Third, overall mission numbers have continuously been increasing with a stable proportion of around 45%, requiring an EMS physician on site [3]. This has led to a higher workload and longer arrival times [4], and moreover, this trend could not be reversed, despite a nationwide increase in EMS physician stations.

Objectives

To overcome these shortcomings, we developed a holistic prehospital telemedical emergency service for the City of Aachen. The routine EMS, including separately dispatchable paramedics and physicians, was complemented with an additional tele-EMS physician [5,6]. During emergency missions for which only paramedics were dispatched, these paramedics can—if needed—either request an on-site EMS physician to be dispatched to the scene or consult with the tele-EMS physician for the treatment of the patient. On consultation, the tele-EMS physician communicates via voice and has immediate access to the vital data of the patient and the Global Positioning System (GPS) coordinates of the ambulance. Paramedics can instantaneously send pictures to the tele-EMS physician, and video can be streamed from the inside of the ambulance. This setting allows for the telemedical delegation (teledelegation) of predefined medications including opioids and has been demonstrated to provide patients with standard care in accordance with treatment guidelines and without complications [7-9]. After initial establishment and evaluation [6,10], the system has been integrated as an around-the-clock routine component of the City of Aachen EMS. In this retrospective analysis of the first 3 operational years, we sought to assess the utilization, safety, and technical performance of the holistic Aachen prehospital telemedical emergency service.

Methods

Patients

All patients in primary emergency missions with consultation of the tele-EMS physician during the first 3 operational years (April 2014 to March 2017) were included in the analyses. Tele-EMS physicians operate from a tele-EMS control center in close proximity to the mission control center of the EMSs. Tele-EMS physicians communicate with especially equipped ambulances via voice and have immediate access to the vital data of the patient and the GPS coordinates of the ambulance. In addition, video can be streamed from the inside of the ambulance. After making a working diagnosis, tele-EMS physicians are prompted with a guideline-based SOP, including corresponding treatment algorithms and corresponding checklist. Previous publications provide a detailed description of the establishment and functionality as well as technical details of the Aachen tele-EMS physician [5-7,11].

Data Sources and Analyses

Mission and Patient Data

Age, sex, National Advisory Committee for Aeronautics (NACA) scores, categorized main symptoms, administration of nonopioid drugs and opioid drugs, and teleconsultation times were abstracted from the protocols archived by the tele-EMS physician. Transport modalities and mission times of involved mobile EMS forces (paramedics and on-site EMS physician) were abstracted from the control system of the dispatch center (COBRA 4; ISE).

Safety

Adverse events were assessed in a 2-step process. First, we preselected protocols if (1) the free-text comments section contained parts of or the German words or word combinations for unsuccessful, instable, unstable, no sign of recovery, on-site physician, hypotensive, hypotension, allergy, allergic, anaphylaxis, anaphylactic, accident, accidental, erroneous, error, confusion, confused, and possible misspellings; or (2) the tele-EMS physician administered catecholamines (adrenaline, noradrenaline, and theodrenaline-cafedrine), antihistamines, or corticosteroids. Second, these protocols were reviewed by two independent researchers, and disagreements were resolved by consensus.

Technical Performance

During these first 3 operational years, tele-EMS physicians were asked to regularly fill out a paper-based questionnaire after each teleconsultation to assess the technical performance of the system. *Transmission quality of the telemetric components*, *two-way voice communication*, *GPS coordinates*, *real-time vital data*, *12-lead restingelectrocardiogram (ECG)*, *still pictures*, and *one-way video stream* were assessed in mutually exclusive categories *no malfunctions*; *some malfunctions, quality not affected*; *malfunctions, quality reduced*; and *malfunctions, transmission impossible*. Teleconsultations with impossible transmission of voice communication, real-time vital data, 12-lead resting ECG, still pictures, and video stream were considered complete system failures.

Statistical Analyses

Frequencies are reported as proportion and percentage and numerical values as median and interquartile range (IQR), without prior testing for normal distribution. Trends were fitted using univariable linear regression. Slopes were reported with 95% CI and whether the slope differed significantly from 0. CIs for the proportion of a count were analyzed using the method described by Wilson [12]. Linear regression modeling was conducted, and figures were created with Prism version 8.0.2 (GraphPad) for Mac operating system (OS). All other analyses were conducted with RStudio version 1.1.463 for Mac OS (RStudio) operating R version 3.5.2 for Mac OS. A type I error of 5% or less was considered statistically significant.

Patient and Public Involvement

Patients and the public were not involved in the design, conduct, and reporting of the research.

Ethics Approval and Consent to Participate

The local ethics committee granted analysis of the data for quality assurance purposes and waived the requirement of informed consent (EK109/15, University Hospital RWTH Aachen).

Availability of Data and Material

The datasets used and/or analyzed during this study are available from the corresponding author on reasonable request.

Results

Patients, Mission, and Patient data

During the study period, tele-EMS physicians were consulted in primary emergency missions for 6265 patients. Patients had a median age of 70 years (IQR 48-81), and 52.88% (3313/6265) were female. Of the 6265 patients, most were categorized as NACA III (3594/6265, 57.36%) and NACA IV (1445/6265, 23.06%) cases. A total of 6.43% presented a life-threatening condition (403/6265), indicated by an NACA score of V or greater. The majority of patients (4328/6265, 69.08%) were treated on-site by the tele-EMS physician and subsequently transferred to a hospital, as little as 7.80% (489/6265) were neither treated on-site nor transferred to the hospital (Table 1).

Table 1. Characteristics of the patients treated in 6265 teleconsultations in primary emergency missions by tele-emergency medical service physicians of the Aachen telemedical prehospital emergency service.

Characteristics	Values
Age (years), median (IQR)[a,b]	70 (48-81)
Sex, n (%)[c]	
Female	3313 (52.88)
Male	2844 (45.40)
Severity of the emergency, n (%)[d]	
NACA[e] I	47 (0.75)
NACA II	457 (7.29)
NACA III	3594 (57.37)
NACA IV	1445 (23.06)
NACA V	394 (6.29)
NACA VI	3 (0.05)
NACA VII	8 (0.13)
Type of main symptom, n (%)	
Circulatory	1146 (18.29)
Neurologic	1049 (16.74)
Cardiac	852 (13.60)
Trauma	600 (9.58)
Abdominal	536 (8.56)
Other	2082 (33.23)
Mission details, n (%)[f]	
On-site treatment and transfer to the hospital	4326 (69.05)
Transfer to the hospital only and no on-site treatment	1438 (22.95)
On-site treatment only and no transfer to the hospital	338 (5.40)
No on-site treatment and no transfer to the hospital	489 (7.81)

[a]IQR: interquartile range.

[b]Data of 232 cases missing.

[c]Data of 107 cases missing.

[d]Data of 317 cases missing

[e]NACA: National Advisory Committee for Aeronautics.

[f]Do not add up to 100% because categories are not mutually exclusive.

Utilization of the System

Teleconsultations increased during the first 4 quarters of routine operations from 152 to 420, as subsequently more ambulances became technically equipped for consultations with the tele-EMS physician (Multimedia Appendix 1). On successful implementation of an around-the-clock tele-EMS physician service for 11 ambulances from the second quarter of 2015, teleconsultations further increased by 25.9 per quarter (95% CI 9.1-42.6; *P*=.009). In total, the tele-EMS physician delegated the application of nonopioid drugs in 81.5% (5111/6265; 95% CI 80.6%-82.5%) and opioid drugs in 20.7% (1297/6265; 95% CI 19.7%-21.7%) of the teleconsultations (Multimedia Appendix 2). The teledelegated application of nonopioids increased similarly with increasing mission numbers (17.3 per quarter; 95% CI 3.1-31.4; *P*=.02), as did the teledelegated application of opioids (12.0 per quarter; 95% CI 0.8-23.3; *P*=.04). In fact, when analyzing all emergency missions involving any EMS physician in the city of Aachen, the proportion of missions carried out with the newly established tele-EMS physician continuously increased, whereas the proportion of missions carried out with a conventional on-site EMS physician continuously decreased (0.9% per quarter; 95% CI 0.3%-1.4%; *P*=.08; Multimedia Appendix 1).

Overall, 91.10% of the teleconsultations (5708/6265; 95% CI 90.3%-91.8%) were used to support the on-site paramedics. In 3.65% of the teleconsultations (229/6265; 95% CI 3.2%-4.2%), the tele-EMS physician saw the need to dispatch an on-site EMS physician. On-site paramedics consulted with the tele-EMS

physician in 2.45% (154/6265; 95% CI 2.1%-2.9%) of the missions to bridge the time until a simultaneously dispatched EMS physician arrived on the scene. The tele-EMS was contacted for support after an on-site physician already had arrived on site in 4.38% (275/6265; 95% CI 3.9%-4.9%) of all teleconsultations.

Safety

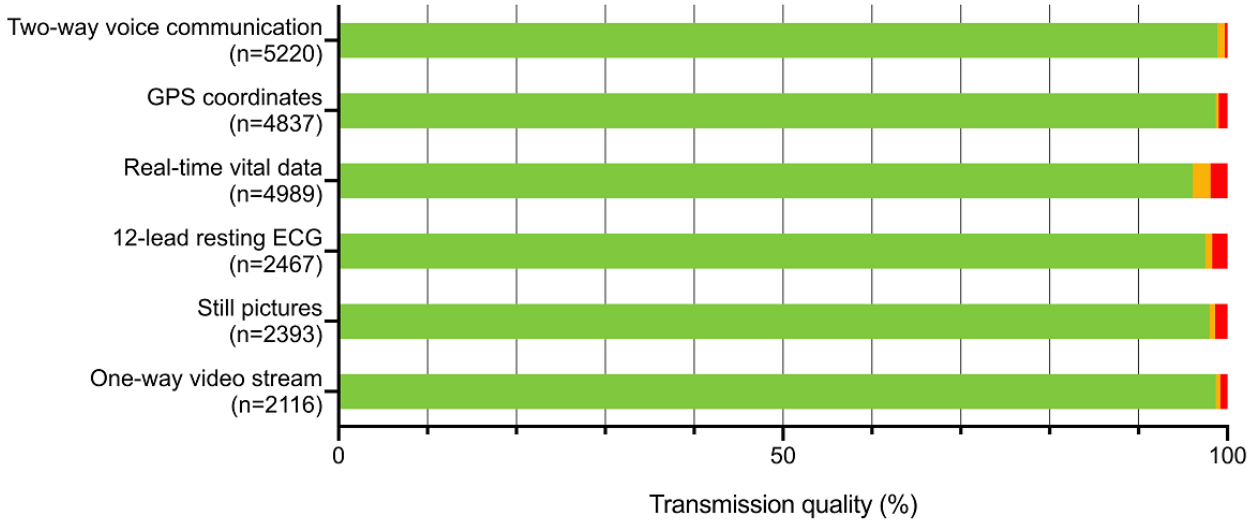
Only 6 of 6265 patients (0.09%; 95% CI 0.04%-0.21%) experienced adverse events. One patient became unconscious after being treated with teledelegated nitroglycerine for acute coronary syndrome, an on-site physician was immediately dispatched to the scene by the tele-EMS physician, and the patient regained consciousness after 10 seconds and presented stable vital signs and full orientation until arrival at the hospital. A 94-year-old woman entered a highly agitated state after receiving ketamine for the treatment of a hip fracture. An on-site physician was dispatched, and the patient was sedated with midazolam and safely transferred to the hospital. One patient received reproterol (Bronchospasmin, a beta-2-mimetic) instead of scopolamine (Buscopan, a parasympatholytic) because of the phonetical similarity of the German trade names; however, the patient did not suffer from any side effects. Paramedics accidentally punctured the cubital artery instead of a vein, which was instantaneously recognized and adequately treated with a pressure bandage. In addition, 2 patients suffered from erythema: 1 patient presented a generalized erythema and pruritus to the application of morphine (sufficiently treated with histamine antagonists), and the other presented a local erythema after the infusion of metamizole (the infusion was stopped, and no signs of generalized anaphylactic reactions were observed). All patients with adverse events presented normal vital data at arrival in the emergency department.

Technical Performance

Tele-EMS physicians assessed the technical performance of the system by filling out the paper-based questionnaire following 5220 of all 6265 teleconsultations (83.32%). Two-way voice communication was used in all 5220 (100.00%) assessed teleconsultations. Real-time vital data were transmitted in 95.57% (4989/5220), GPS coordinates in 92.66% (4837/5220), 12-lead resting ECGs in 47.26% (2467/5220), still pictures in 45.84% (2393/5220), and one-way video streams in 44.25% (2310/5220) of consultations with the tele-EMS physician.

Complete malfunctions of single transmission components with no possible transmission occurred from as low as 0.26% (14/5220; 95% CI 0.2%-0.5%) during two-way voice communications, more than 1.01% (53/5220; 95% CI 0.8%-1.3%) during GPS coordinate transmissions, 1.34% (70/5220; 95% CI 1.1%-1.7%) during one-way video stream transmissions, 1.43% (72/5220; 95% CI 1.1%-1.7%) during still picture transmissions, 1.66% (87/5220; 95% CI 1.3%-2.1%) during 12-lead resting ECG transmissions to as high as 1.91% (100/5220; 95% CI 1.6%-2.3%) during real-time vital data transmissions (Figure 1). Complete system failures occurred in only 0.34% (18/5220; 95% CI 0.2%-0.6%) of all teleconsultations.

Figure 1. Technical performance of the Aachen physician-staffed telemedical prehospital emergency service in 5220 of 6265 teleconsultations by transmission component and quality: transmission quality not affected (green), transmissions with reduced quality (yellow), and complete transmission component malfunction (red). ECG: electrocardiogram; GPS: Global Positioning System.



Discussion

Principal Findings

In response to the increasing mission numbers involving on-site physicians in German EMSs, we had developed a holistic prehospital telemedical emergency service for the City of Aachen, which began routine operations in 2014. In this retrospective analysis of the first 3 operational years, we sought to assess the utilization, safety, and technical performance of the system. We found that paramedics consulted the tele-EMS physician on a regular basis in 6265 primary emergency missions with a continuing increase of 26 teleconsultations per quarter. Of all missions involving any EMS physician, the proportion of missions carried out with the tele-EMS physician increased by 4% per year. Telemedical treatment by the tele-EMS physician resulted in as little as 6 complications in 6265 teleconsultations (0.1%). In a convenience sample

assessing the technical performance of the system, we found that single transmission components malfunctioned within a range of 0.3% (two-way voice communication) to 1.9% (real-time vital data stream) in 5220 teleconsultations. Complete system failures occurred very rarely in only 18 teleconsultations (0.3%).

Strengths and Weaknesses and Comparison With Other Studies

This is the world's largest analysis of cases that were provided with routine medical care within a holistic telemedical system for prehospital emergency care. We were able to analyze all teleconsultations for primary emergency missions during the first 3 operational years of the system. The fact that we did not analyze a sample combined with the high number of cases contributes to the internal validity of our results. On the other hand, the uniqueness of the Aachen prehospital telemedical emergency service limits the generalizability to other (even similar) systems.

The steep increase in teleconsultations during the first operational year has to be ascribed primarily to the increasing number of ambulances (from 2 to 11) being subsequently equipped for telemedical operations. In addition, the tele-EMS service expanded its availability from a 12-hour day to an around-the-clock service during the first year. However, during the 2-year routine service with an around-the-clock tele-EMS physician and 11 operational ambulances, the number of teleconsultations further increased, as did the teledelegation of nonopioid drugs and opioids. This is most likely because of a training effect operating the new telemedical emergency care system. Paramedics may have overcome barriers in applying the new system more easily [13] and may have recognized increasing possibilities to work with the new system instead of requesting a *traditional* on-site EMS physician [14,15].

The number of complications in our analyses was low. This may have been because of the fact that we analyzed routine treatment protocols and that complications or adverse events may not have been documented. Complications such as the occurrence of common side effects (nausea after the administration of opioids and relevant hypotension after the administration of metamizole), mild allergic reactions, or confusion of phonetically similar drugs (eg, midazolam and metamizole) may have occurred but not explicitly mentioned as they presumably could also have occurred with EMS physician present on site. Moreover, as all our tele-EMS physicians are anesthesiologists (board certified for anesthesiology and prehospital emergency care, longstanding experience in more than 500 on-site prehospital emergency missions, and certified provider of advanced life support and prehospital trauma life support), they could have easily considered these events as typical and treatable. In addition, none of the recorded complications were inherent in the system and could have easily happened exactly similar to this with an EMS physician treating the patient on-site. Still, a low complication rate concurs with other emergency telemedicine studies that have demonstrated high safety in comparable *nonserious* emergency patient populations groups [16,17] but also for patients with life-threatening conditions as, for example, ST-elevation myocardial infarction [18].

High rates of successful data transmission in our analysis concur with or outperform other studies that demonstrated telemedical data transmissions from mobile emergency units to be safe and reliable [11,19-21]. Being the most redundantly secured transmission component (headset from the Global System for Mobile Communications–equipped vital sign monitor, backup mobile phone, and any landline can be used), it comes as no surprise that two-way voice communication was least affected by transmission malfunctions. In our sample, a complete system failure without possible transmission occurred in only 0.3% of the teleconsultations. These events are highly likely caused by bad mobile service reception and can often easily be solved by relocating the transmission unit, that is, the ambulance. However, we are unable to determine whether technical problems were caused by malfunctions in the general telecommunication infrastructure or the telemedical service itself.

Unfortunately, the technical performance questionnaire was not filled out in 17% of the missions, leaving 83% of all teleconsultations to be analyzed. As this may very well have introduced bias to our study, this has most likely led to an overestimation of technical malfunctions, as tele-EMS physicians would rather fill out the report form after having experienced technical difficulties during the mission. In addition, we reviewed the mission protocols of the cases with complete transmission failures and found no irregularities, the missions were completed without adverse events, and the patients transferred to a hospital. The highest failure rate was detected when transmitting real-time vital data. This malfunction—and with the exception of 12-lead resting ECG and all other transmission components—can temporarily be overcome if the voice communication between the tele-EMS and the paramedics is still intact. In this manner, even the backup mobile phone's automatic transmission of still pictures can be used as a fallback solution for failing 12-lead resting ECG transmissions.

Besides being frequently used, safe, and technically feasible, the holistic Aachen prehospital emergency system has been demonstrated to be effective, of high quality, and to provide instantaneous care to patients [22-26]: we have provided evidence in the past that the Aachen tele-EMS physician is as effective as conventional on-site EMS physicians when treating patients with pain [6,10], stroke [7], hypertensive crises [9], and acute coronary syndrome [8]. Similar effects have been provided by other work groups, as, for example, the prehospital telemedicine emergency care for patients with stroke leads to faster diagnosis and faster treatment compared with standard care [27,28]. These effects are most likely to be even more pronounced in rural areas where paramedics are noticeably faster available compared with on-site EMS physicians [29,30].

Policy Implications and Conclusions

Life-threatening complications and errors can also occur in prehospital emergency care provided by an on-site physician [31]. During in-hospital life-threatening emergencies, it is possible and common for physicians to seek instantaneous support by more experienced colleagues. The on-site EMS

physician does not have this possibility. Surprisingly, 275 teleconsultations in our study were made after an on-site physician arrived on the scene, presumably to provide support for the on-site EMS physician. The analyses of these cases will provide more insight into the effective application of our system for the purpose of on-site EMS physician support.

The increasing use of checklists in prehospital emergency care has fostered the adherence to standards and overall treatment quality [32]. The Aachen holistic prehospital telemedicine emergency service was designed to foster guideline-based treatment of patients. For this purpose, on entering a working diagnosis, the operating tele-EMS physician is automatically supplied with a guideline-based SOP, including treatment algorithms and corresponding checklists [6-9,11]. It is highly likely that the quality of care and safety of such a system is at least comparable with standard care provided by on-site EMS physicians for non–life-threatening emergencies. Therefore, this hypothesis is currently tested in a large randomized controlled trial [33].

In conclusion, this retrospective analysis of primary emergency care teleconsultations conducted during the first 3 operational years showed the Aachen prehospital EMS to be a frequently used, safe, and technically reliable system to provide medical care for prehospital emergency patients, without an EMS physician present. This telemedical emergency service is likely to provide high-quality emergency care for non–life-threatening events and could relieve mobile EMS physicians' workload. Consequently, this may free EMS physicians for the treatment of more serious, life-threatening emergency missions and therefore provide faster high-quality care for all patients treated in EMSs.

Acknowledgments

The authors thank Janine Eßer, Klaudia Lucyna Ogrodzki, and Carina Pabst for their help in assessing the complication rates.

Authors' Contributions

All authors analyzed and interpreted the data equally. MF and RR were main contributors to manuscript writing. FK was the main contributor to data analysis. All authors read and approved the final manuscript.

Conflicts of Interest

MF, SKB, FK, and AS declare no conflicts of interest. JCB, MC, and RR are shareholders of the Docs-In-Clouds company Aachen, Germany. FH is employee of the P3 telehealthcare company Aachen, Germany.

Multimedia Appendix 1

Utilization of the Aachen prehospital telemedicine emergency system during the first 3 operational years. (A) Consultations of the tele-EMS physician for primary emergency missions (dark blue) increased strongly during the run-in phase (hatched gray), as did the teledelegated applications of nonopioid drugs (light blue) and opioids (orange). The number of the missions increased further after the tele-EMS physician became available around the clock for 11 ambulances at the beginning of the second quarter in 2015. (B) Of all primary emergency missions carried out including an EMS physician, the proportion of missions carried out with a tele-EMS physician increased by 4% ($P < .05$) after the run-in phase. EMS: emergency medical service.
[PDF File (Adobe PDF File), 213 KB - jmir_v21i10e14907_app1.pdf]

Multimedia Appendix 2

Missions with on-site and tele-emergency medical service physicians in the City of Aachen during the first three operational years.
[PDF File (Adobe PDF File), 27 KB - jmir_v21i10e14907_app2.pdf]

References

1. Luiz T, van Lengen RH, Wickenkamp A, Kranz T, Madler C. [Operational availability of ground-based emergency medical services in Rheinland-Palatinate: state-wide web-based system for collation, display and analysis]. Anaesthesist 2011 May;60(5):421-426. [doi: 10.1007/s00101-010-1826-3] [Medline: 21184039]
2. Schuster M, Pints M, Fiege M. Duration of mission time in prehospital emergency medicine: effects of emergency severity and physicians level of education. Emerg Med J 2010 May;27(5):398-403. [doi: 10.1136/emj.2009.074211] [Medline: 20442179]
3. Bundestag D. DIP21. 2014. Bericht Über Maßnahmen Auf Dem Gebiet Der Unfallverhütung Im Straßenverkehr 2012 Und 2013: Unfallverhütungsbericht Straßenverkehr 2012/2013 URL: http://dip21.bundestag.de/dip21/btd/18/024/1802420.pdf [accessed 2018-03-07]
4. Joó S. Deutsches Ärzteblatt. 2000. Rettungsdienst: Starke Leistungsbilanz URL: https://www.aerzteblatt.de/archiv/25088/Rettungsdienst-Starke-Leistungsbilanz [accessed 2019-08-23]

5. Brokmann JC, Rossaint R, Bergrath S, Valentin B, Beckers SK, Hirsch F, et al. [Potential and effectiveness of a telemedical rescue assistance system. Prospective observational study on implementation in emergency medicine]. Anaesthesist 2015 Jun;64(6):438-445. [doi: 10.1007/s00101-015-0039-1] [Medline: 26036316]
6. Brokmann JC, Rossaint R, Hirsch F, Beckers SK, Czaplik M, Chowanetz M, et al. Analgesia by telemedically supported paramedics compared with physician-administered analgesia: a prospective, interventional, multicentre trial. Eur J Pain 2016 Aug;20(7):1176-1184. [doi: 10.1002/ejp.843] [Medline: 26914284]
7. Bergrath S, Reich A, Rossaint R, Rörtgen D, Gerber J, Fischermann H, et al. Feasibility of prehospital teleconsultation in acute stroke--a pilot study in clinical routine. PLoS One 2012;7(5):e36796 [FREE Full text] [doi: 10.1371/journal.pone.0036796] [Medline: 22629331]
8. Brokmann JC, Conrad C, Rossaint R, Bergrath S, Beckers SK, Tamm M, et al. Treatment of acute coronary syndrome by telemedically supported paramedics compared with physician-based treatment: a prospective, interventional, multicenter trial. J Med Internet Res 2016 Dec 1;18(12):e314 [FREE Full text] [doi: 10.2196/jmir.6358] [Medline: 27908843]
9. Brokmann JC, Rossaint R, Müller M, Fitzner C, Villa L, Beckers SK, et al. Blood pressure management and guideline adherence in hypertensive emergencies and urgencies: a comparison between telemedically supported and conventional out-of-hospital care. J Clin Hypertens (Greenwich) 2017 Jul;19(7):704-712 [FREE Full text] [doi: 10.1111/jch.13026] [Medline: 28560799]
10. Lenssen N, Krockauer A, Beckers SK, Rossaint R, Hirsch F, Brokmann JC, et al. Quality of analgesia in physician-operated telemedical prehospital emergency care is comparable to physician-based prehospital care - a retrospective longitudinal study. Sci Rep 2017 May 8;7(1):1536 [FREE Full text] [doi: 10.1038/s41598-017-01437-5] [Medline: 28484212]
11. Bergrath S, Rörtgen D, Rossaint R, Beckers SK, Fischermann H, Brokmann JC, et al. Technical and organisational feasibility of a multifunctional telemedicine system in an emergency medical service - an observational study. J Telemed Telecare 2011;17(7):371-377. [doi: 10.1258/jtt.2011.110203] [Medline: 21933897]
12. Wilson EB. Probable inference, the law of succession, and statistical inference. J Am Stat Assoc 1927 Jun;22(158):209-212. [doi: 10.2307/2276774]
13. Sjögren LH, Törnqvist H, Schwieler A, Karlsson L. The potential of telemedicine: barriers, incentives and possibilities in the implementation phase. J Telemed Telecare 2001;7(Suppl 1):12-13. [doi: 10.1177/1357633X010070S105] [Medline: 11576475]
14. Granger D, Vandelanotte C, Duncan MJ, Alley S, Schoeppe S, Short C, et al. Is preference for mhealth intervention delivery platform associated with delivery platform familiarity? BMC Public Health 2016 Jul 22;16:619 [FREE Full text] [doi: 10.1186/s12889-016-3316-2] [Medline: 27450240]
15. Molfenter T, Boyle M, Holloway D, Zwick J. Trends in telemedicine use in addiction treatment. Addict Sci Clin Pract 2015 May 28;10:14 [FREE Full text] [doi: 10.1186/s13722-015-0035-4] [Medline: 26016484]
16. Dale J, Williams S, Foster T, Higgins J, Snooks H, Crouch R, et al. Safety of telephone consultation for 'non-serious' emergency ambulance service patients. Qual Saf Health Care 2004 Oct;13(5):363-373 [FREE Full text] [doi: 10.1136/qhc.13.5.363] [Medline: 15465940]
17. Benger JR, Noble SM, Coast J, Kendall JM. The safety and effectiveness of minor injuries telemedicine. Emerg Med J 2004 Jul;21(4):438-445 [FREE Full text] [doi: 10.1136/emj.2002.004028] [Medline: 15208226]
18. Keeling P, Hughes D, Price L, Shaw S, Barton A. Safety and feasibility of prehospital thrombolysis carried out by paramedics. Br Med J 2003 Jul 5;327(7405):27-28 [FREE Full text] [doi: 10.1136/bmj.327.7405.27] [Medline: 12842952]
19. Yperzeele L, van Hooff RJ, de Smedt A, Espinoza AV, van Dyck R, van de Casseye R, et al. Feasibility of AmbulanCe-Based Telemedicine (FACT) study: safety, feasibility and reliability of third generation in-ambulance telemedicine. PLoS One 2014;9(10):e110043 [FREE Full text] [doi: 10.1371/journal.pone.0110043] [Medline: 25343246]
20. Cabrera MF, Arredondo MT, Quiroga J. Integration of telemedicine into emergency medical services. J Telemed Telecare 2002;8(Suppl 2):12-14. [doi: 10.1177/1357633X020080S206] [Medline: 12217117]
21. Felzen M, Brokmann JC, Beckers SK, Czaplik M, Hirsch F, Tamm M, et al. Improved technical performance of a multifunctional prehospital telemedicine system between the research phase and the routine use phase - an observational study. J Telemed Telecare 2017 Apr;23(3):402-409. [doi: 10.1177/1357633X16644115] [Medline: 27080747]
22. Audebert HJ, Kukla C, von Claranau SC, Kühn J, Vatankhah B, Schenkel J, TEMPiS Group. Telemedicine for safe and extended use of thrombolysis in stroke: the telemedic pilot project for integrative stroke care (TEMPiS) in Bavaria. Stroke 2005 Feb;36(2):287-291. [doi: 10.1161/01.STR.0000153015.57892.66] [Medline: 15625294]
23. Audebert HJ, Schenkel J, Heuschmann PU, Bogdahn U, Haberl RL, Telemedic Pilot Project for Integrative Stroke Care Group. Effects of the implementation of a telemedical stroke network: the telemedic pilot project for integrative stroke care (TEMPiS) in Bavaria, Germany. Lancet Neurol 2006 Sep;5(9):742-748. [doi: 10.1016/S1474-4422(06)70527-0] [Medline: 16914402]
24. Chen MJ, Chen KY, Chiang SJ, Daimon M, Lee JS, Yu EW, et al. A telehealth service model for the treatment of hypertension. J Telemed Telecare 2013 Jul;19(5):238-241. [doi: 10.1177/1357633X13495476] [Medline: 24163232]
25. Demaerschalk BM, Raman R, Ernstrom K, Meyer BC. Efficacy of telemedicine for stroke: pooled analysis of the stroke team remote evaluation using a digital observation camera (STRokE DOC) and STRokE DOC Arizona telestroke trials. Telemed J E Health 2012 Apr;18(3):230-237 [FREE Full text] [doi: 10.1089/tmj.2011.0116] [Medline: 22400970]

26. Herrington G, Zardins Y, Hamilton A. A pilot trial of emergency telemedicine in regional western Australia. J Telemed Telecare 2013 Oct;19(7):430-433. [doi: 10.1177/1357633X13506531] [Medline: 24218359]
27. Taqui A, Cerejo R, Itrat A, Briggs FB, Reimer AP, Winners S, Cleveland Pre-Hospital Acute Stroke Treatment (PHAST) Group. Reduction in time to treatment in prehospital telemedicine evaluation and thrombolysis. Neurology 2017 Apr 4;88(14):1305-1312. [doi: 10.1212/WNL.0000000000003786] [Medline: 28275084]
28. Mohr NM, Young T, Harland KK, Skow B, Wittrock A, Bell A, et al. Telemedicine is associated with faster diagnostic imaging in stroke patients: a cohort study. Telemed J E Health 2019 Feb;25(2):93-100. [doi: 10.1089/tmj.2018.0013] [Medline: 29958087]
29. Mohr NM, Young T, Harland KK, Skow B, Wittrock A, Bell A, et al. Emergency department telemedicine shortens rural time-to-provider and emergency department transfer times. Telemed J E Health 2018 Aug;24(8):582-593. [doi: 10.1089/tmj.2017.0262] [Medline: 29293413]
30. Brunetti ND, di Pietro G, Aquilino A, Bruno AI, Dellegrottaglie G, di Giuseppe G, et al. Pre-hospital electrocardiogram triage with tele-cardiology support is associated with shorter time-to-balloon and higher rates of timely reperfusion even in rural areas: data from the Bari- Barletta/Andria/Trani public emergency medical service 118 registry on primary angioplasty in ST-elevation myocardial infarction. Eur Heart J Acute Cardiovasc Care 2014 Sep;3(3):204-213. [doi: 10.1177/2048872614527009] [Medline: 24604713]
31. Timmermann A, Russo SG, Eich C, Roessler M, Braun U, Rosenblatt WH, et al. The out-of-hospital esophageal and endobronchial intubations performed by emergency physicians. Anesth Analg 2007 Mar;104(3):619-623. [doi: 10.1213/01.ane.0000253523.80050.e9] [Medline: 17312220]
32. Kerner T, Schmidbauer W, Tietz M, Marung H, Genzwuerker HV. Use of checklists improves the quality and safety of prehospital emergency care. Eur J Emerg Med 2017 Apr;24(2):114-119. [doi: 10.1097/MEJ.0000000000000315] [Medline: 26287802]
33. Stevanovic A, Beckers SK, Czaplik M, Bergrath S, Coburn M, Brokmann JC, TEMS Collaboration Group. Telemedical support for prehospital emergency medical service (TEMS trial): study protocol for a randomized controlled trial. Trials 2017 Jan 26;18(1):43 [FREE Full text] [doi: 10.1186/s13063-017-1781-2] [Medline: 28126019]

Abbreviations

ECG: electrocardiogram
EMS: emergency medical service
GPS: Global Positioning System
IQR: interquartile range
NACA: National Advisory Committee for Aeronautics
OS: operating system
SOP: standard operating procedure

Edited by G Eysenbach; submitted 02.06.19; peer-reviewed by M Bernhard, L Ranandeh, M Gaman, J Farzi; comments to author 25.06.19; revised version received 19.07.19; accepted 21.07.19; published 08.10.19.

Please cite as:
Felzen M, Beckers SK, Kork F, Hirsch F, Bergrath S, Sommer A, Brokmann JC, Czaplik M, Rossaint R
Utilization, Safety, and Technical Performance of a Telemedicine System for Prehospital Emergency Care: Observational Study
J Med Internet Res 2019;21(10):e14907
URL: https://www.jmir.org/2019/10/e14907
doi:10.2196/14907
PMID:31596244

Review

Reaching People With Disabilities in Underserved Areas Through Digital Interventions: Systematic Review

Leming Zhou[1], PhD; Bambang Parmanto[1], PhD

Department of Health Information Management, University of Pittsburgh, Pittsburgh, PA, United States

Corresponding Author:
Leming Zhou, PhD
Department of Health Information Management
University of Pittsburgh
3600 Forbes Ave at Meyran Ave
6021 Forbes Tower
Pittsburgh, PA, 15260
United States
Phone: 1 4123836653
Email: Leming.Zhou@pitt.edu

Abstract

Background: People with disabilities need rehabilitation interventions to improve their physical functioning, mental status, and quality of life. Many rehabilitation interventions can be delivered electronically ("digitally") via telehealth systems. For people with disabilities in underserved areas, electronically delivered rehabilitation interventions may be the only feasible service available for them.

Objective: The objective of this study was to evaluate the current status of digital interventions for people with disabilities in remote and underserved areas.

Methods: A systematic review was conducted on this topic. Keyword searches in multiple databases (PubMed, CINAHL, and Inspec) were performed to collect articles published in this field. The obtained articles were selected based on our selection criteria. Of the 198 identified articles, 16 duplicates were removed. After a review of the titles and abstracts of the remaining articles, 165 were determined to be irrelevant to this study and were therefore removed. The full texts of the remaining 17 articles were reviewed, and 6 of these articles were removed as being irrelevant to this study. The 11 articles remaining were discussed and summarized by 2 reviewers.

Results: These 11 studies cover a few types of disabilities, such as developmental disabilities and mobility impairments as well as several types of disability-causing disorders such as stroke, multiple sclerosis, traumatic brain injury, and facio-scapulo-humeral muscular dystrophy. Most of these studies were small-scale case studies and relatively larger-scale cohort studies; the project evaluation methods were mainly pre-post comparison, questionnaires, and interviews. A few studies also performed objective assessment of functional improvement. The intervention technology was mainly videoconferencing. Moreover, 10 of these studies were for people with disabilities in rural areas and 1 was for people in urban communities.

Conclusions: A small number of small-scale studies have been conducted on digital interventions for people with disabilities in underserved areas. Although the results reported in these studies were mostly positive, they are not sufficient to prove the effectiveness of telehealth-based digital intervention in improving the situation among people with disabilities because of the small sample sizes and lack of randomized controlled trials.

(J Med Internet Res 2019;21(10):e12981) doi:10.2196/12981

KEYWORDS

systematic review; digital intervention; electronic intervention; e-intervention; underserved area; disability; telemedicine; telerehabilitation; eHealth; digital health

Introduction

Background

Advancement in health care technologies has contributed to the decline of mortality in the United States in recent years. Many people survive diseases that used to have a high mortality rate [1]. However, surviving a severe disease is often not the end of the story but the beginning of a life filled with many serious challenges. We use 2 specific examples to demonstrate these challenges.

In the United States, although stroke is still one of the top 10 leading causes of death [2] and stroke prevalence has increased in recent years because of the increase in the aging population [3], the actual number of deaths caused by stroke declined by 2.3% from 2005 to 2015 [1]. There are approximately 6.8 million stroke survivors in the United States [1], but only 10% of stroke survivors recover almost completely. Nearly 65% of stroke patients experience significant or permanent disabilities such as paralysis, urinary incontinence, aphasia, and cognitive disability. In all, about 3% of adults in the United States claim that they were disabled because of stroke [1]. These disabled stroke survivors need long-term rehabilitation services to relearn the skills lost in the stroke attack. Previous research has reported that people who participate in focused poststroke rehabilitation programs perform better than most people who do not undergo poststroke rehabilitation [4-6]. Depending on the complications after stroke, the rehabilitation plan for each patient must be personalized, and the stroke rehabilitation may involve a variety of specialists for a long period.

Traumatic brain injury (TBI) is another leading cause of death and long-term disability. It is estimated that 1.7 million TBIs occur in the United States annually, which contribute to approximately 30% of all injury deaths. Advances in trauma care have resulted in increases in the number of survivors of TBI in recent years. In the United States, there are approximately 5.3 million individuals living with disabilities caused by TBI [7]. For TBI survivors, there may be physical, behavioral, and psychological alterations that require extensive rehabilitation services over a long period to restore skill [7,8].

There are also other causes of disability that may not be as severe as stroke or TBI but which also present similar challenges. Multiple sclerosis (MS) is a neurological disorder that leads to several conditions and disabilities that limit several daily life activities [9,10]. Some of these conditions are associated with low rates of physical activity, including fatigue, weakness, falls, and depression [11,12]. Previous studies have demonstrated that exercise and physical activity can reduce symptoms, increase physical health, and improve quality of life for people with MS [13].

Developmental disabilities (DDs) are a group of conditions such as intellectual disability, learning disability, cerebral palsy (CP), hearing loss, autism spectrum disorder, attention deficit hyperactivity disorder, and other developmental delays because of physical, language, learning, or behavior impairments that happened as an infant or during development as a fetus. These conditions may impact children's daily functioning and usually last for a lifetime [14]. Recent estimates suggest that about 15% of children aged between 3 and 17 years have one or more development disability (DD) in the United States [15]. Early intervention (EI) is the service and support to children with DD and their families, for example, speech and language therapy (SLT), physical therapy (PT), and other types of services, based on the conditions and needs of the children and their family. EI can have a significant impact on a child's ability to learn new skills and can increase success in life and school. However, EI is more effective if delivered in the natural environment of babies and from the very early stages of childhood, which is challenging for families who live in rural communities.

For many survivors of stroke or TBI and people with MS or DD, the common need is long-term intervention. A comprehensive, coordinated rehabilitation program can reduce secondary complications and improve functional outcomes. Many rehabilitation protocols require multidisciplinary, high-intensity therapy sessions multiple times a week, for weeks, months, or even years [16]. Some rehabilitation services need to be delivered at a very specific time and frequency. All of these requirements are difficult to meet for people in the rural and remote areas as there are extraordinary physical, financial, and logistical hardships.

Specifically, according to the US Census Bureau, approximately 20% of the US population lives in rural areas [17]; however, less than 8% of the nation's physicians are practicing in rural areas. The majority of these physicians are in primary care such as family practice, general internal medicine, and pediatrics [18,19]. In other words, receiving regular health care services in rural areas is challenging, it is even harder for people with disabilities in rural areas to access highly qualified specialists for poststroke, post-TBI, or EI rehabilitation services [18,20-22]. Traveling to major cities and seeking the desired intervention costs a lot in terms of money and time, which can be a very heavy burden for family members of the patients.

People with disabilities in poor urban communities face different challenges in terms of receiving such services. Although geographical distance may not be such a huge issue in terms of accessing health care services, difficulties with transportation, dependence on caregivers, low health literacy, and lower socioeconomic status still create significant challenges in terms of access to high-quality health care for people with disabilities in poor urban communities [23]. For this reason, in this study, both rural areas and poor urban communities are referred to as *underserved areas*.

Digital Interventions

Telehealth may be a viable approach for the delivery of interventions to people with disabilities in underserved areas [24]. The concept of telehealth has been discussed since the 1960s, but then it consisted mainly of using telephones to provide communication between patients and health care providers. Since the 1990s, the availability of the internet has made it possible to use multiple information and communication technologies (ICTs) to deliver digital interventions via telehealth. Further development of ICT and the high penetration of broadband connection at home in the 2000s have made it possible for people with disabilities and health care providers

to communicate more easily via videoconferencing (VC). A national survey in June 2019 indicated that 90% of American adults used the internet, 73% of American adults used broadband connections at home, and 63% of rural American homes are connected to the internet, and 17% of US adults only use smartphones to access the internet [25]. According to a national survey performed by Pew Research Center, also in June 2019, the adoption of smartphone and other smart mobile devices has increased dramatically as well in recent years. In 2011, 35% Americans owned smartphones, whereas in 2019, the rate of smartphone ownership is 81% overall. In rural areas, the smartphone ownership rate is 71% [26]. In other words, the improvement of ICT in recent years and the penetration of the internet and mobile devices make it easier than ever to conduct VC, which is the foundation of many telehealth systems. This situation makes telehealth a feasible approach for delivering digital intervention to people with disabilities in underserved areas.

Telehealth has been used to provide assistive technology assessment [27,28], diagnostic evaluations [29,30], assessment and therapy services [31], and consultation opportunities for practitioners and people with disabilities in rural communities [32]. The benefits of telehealth include access to high-quality care, reduced travel time and costs, and increased collaboration among health care providers [29]. Previous studies have indicated that telehealth is a potentially efficient and effective alternative to hospital-based care to deliver patient-satisfying health care services [33,34].

Telehealth enables therapists to deliver rehabilitative services to patients who cannot access health care providers because of physical, financial, and logistical barriers [35]. In recent years, as technologies have become ubiquitous and costs have declined, it has become easier to support the use of telehealth. As a result, research on telehealth has begun to switch from pilot and case studies to validity and reliability of interventions delivered via telehealth.

There are many ways to categorize telehealth services. In this study, we categorize them into three groups based on the specific technologies used in intervention delivery:

1. Regular phone calls, short message service text messages, interactive voice responses (IVRs), and emails
2. VC using technologies such as Skype and video phone
3. Mobile health apps in telehealth practice

Phone calls, text messages, IVRs, and emails can be useful for encouragement and reminders if patients and caregivers are already familiar with the procedure of the intervention. If that is not the case, these methods cannot deliver new and personalized interventions to people with disabilities in underserved areas. Therefore, in this review, we will not include studies using the technologies in the first category.

Objectives

In this study, our goal was to determine the current status of digital interventions delivered to people with disabilities via telehealth in underserved areas; more specifically, we investigated the type of disabilities covered in recent research studies, the number of people with disabilities involved in those studies, the technologies used in the studies, and the outcomes of those studies.

Although a few similar systematic reviews have been conducted in previous years, the covered studies were not high in quality and the results were not generalizable. In this study, we want to determine whether the situation has improved in recent years in terms of telehealth research on those in underserved area.

Methods

Overview

This systematic review was conducted and reported according to the Preferred Reporting Items for Systematic Review and Meta-Analysis guidelines [36]. Methods of the review process and eligibility criteria were established in advance, and the preliminary results were presented orally at the Rehabilitation Engineering and Assistive Technology Society of North America annual conference in 2018.

Literature Search

The keywords used in this study were "(Telehealth OR mHealth OR telerehabilitation OR eHealth OR telemedicine) AND (disability or impairment) AND (underserved OR rural)." In June 2018, the keyword searches were first performed only in PubMed to obtain a general idea of the number of studies in this area. When only "telehealth" was used in the keyword search, there were 27,900 results from PubMed. When "telehealth AND disability" was used, the returned number of studies was 422. When "telehealth AND disability AND (underserved OR rural)" was used, the obtained number of studies was only 64. The numbers of results were similar when other similar keyword combinations were used in PubMed.

In June 2018, the keyword searches were performed for peer-reviewed journal and conference research articles in 3 bibliographic databases: PubMed, CINAHL, and Inspec, without any year restriction. In total, there were 198 articles obtained from the 3 databases using the keyword searches. Moreover, 16 articles were determined to be duplicates and, therefore, were removed from the article list immediately. The studies described in this systematic review were selected from the remaining 182 articles according to the selection criteria given below.

Selection Criteria

Publication Year

During the literature search using keywords, there was no limit on the year of the publication. However, as digital interventions via telehealth became widely available only after 2000, the study purpose itself limited the publication period to between 2000 and 2018. There was no limit on the age of the patients. The language of the selected articles had to be English. The articles had to be research papers published at conferences or in journals. Reviews, abstracts, editorials, workshop summaries, perspectives, opinions, diagnosis methods, and study protocols were excluded. The study could have been performed in any country.

Population

The population was patients with disabilities (eg, developmental, cognitive, vision, intellectual, and mobility impairments, as well as impairments caused by problems such as TBI, stroke, autism, spinal cord injury, CP, MS, and spina bifida) in underserved areas who participated in telehealth-based digital intervention studies before June 2018. Studies about health care providers who received training or offered teleconsultation were not included in this study.

Intervention

The intervention had to be delivered *digitally* via ICTs such as VC-based intervention on mobility for people who had experienced acute stroke. Other examples of interventions are speech language therapy for improving patients' language ability, occupational therapy (OT) for enhancing participation ability, and psychotherapy for managing stress or depression. If the intervention was delivered only via regular telephone without any video component, email, or IVR, the article was removed from this study. If the article was purely about a telehealth IT system development or a patient condition assessment or monitoring (no intervention), it was removed as well. The setting of the intervention could be a home, nursing home, or clinic in an underserved area.

Comparator

The comparator could be either face-to-face intervention or any other control intervention. Articles with no comparison were also included. After all, the purpose of this study was not to determine whether telehealth-based digital interventions are as good as face-to-face interventions but to determine the current status of delivering telehealth-based digital interventions to people with disabilities in underserved areas.

Outcomes

The outcomes of studies that were considered were as follows: participants' satisfaction with the digital intervention, functional improvement in physical and mental aspects, travel time and cost, and general quality of life improvement.

Study Design

The eligible study designs were quantitative, qualitative, and mixed-method studies that explored the outcomes of a telehealth-based digital intervention. Case studies and pilot studies were included because it was possible that they could enable us to understand the change in status over time.

Study Selection

EndNote X7 (Clarivate Analytics) was utilized to manage the articles and collect data from these articles. The selection of the studies was conducted in 3 rounds. In the first round, duplicates were removed from the study list. In the second round, 2 reviewers (LZ and BP) independently reviewed titles and abstracts against the selection criteria, and disagreements were resolved via multiple discussions. In the third round, both reviewers went through the full texts of the remaining articles and made further selection according to the selection criteria.

Quality Assessment

It was important to assess the quality of the selected studies. The quality criteria were used to verify that the selected studies are relevant to this study and the selected studies themselves were methodologically solid [37]. The 11 selected studies were evaluated with regard to the study purpose, literature review, methodology, results obtained, risk of biases in terms of sampling, measurement and intervention, and the conclusion. For this purpose, the quality of these 11 selected studies was evaluated using a modified version of the critical review form created by researchers at McMaster University [38,39]. More specifically, the 10 questions in Textbox 1 were used to evaluate the quality of these studies. If the answer to a question was yes, the score was 1, otherwise the score was 0. Therefore, the maximum quality score that each study could obtain was 10. If the quality score of a study was less than 6, it was removed from the study.

Textbox 1. Questions used for quality assessment selected from the McMaster University critical review form for quantitative studies.

To perform quality assessment on papers, the following 10 questions can be used:
1. Was the purpose stated clearly?
2. Was relevant background literature reviewed?
3a. Was the sample described in detail?
3b. Was sample size justified?
4. Were the outcome measures reliable?
5. Intervention was described in detail?
6a. Results were reported in terms of statistical significance?
6b. Were the analysis methods appropriate?
6c. Clinical importance was reported?
7. Conclusions were appropriate given study methods and results?

Data Extraction and Synthesis

Two reviewers (LZ and BP) extracted data from the 11 articles that were found to meet the inclusion criteria. A standardized form was used for data extraction. Data items on the extraction form include the following: first author's name; publication year; journal or conference name; disability or disability-causing disease; underserved area; sample size; participants' age, sex,

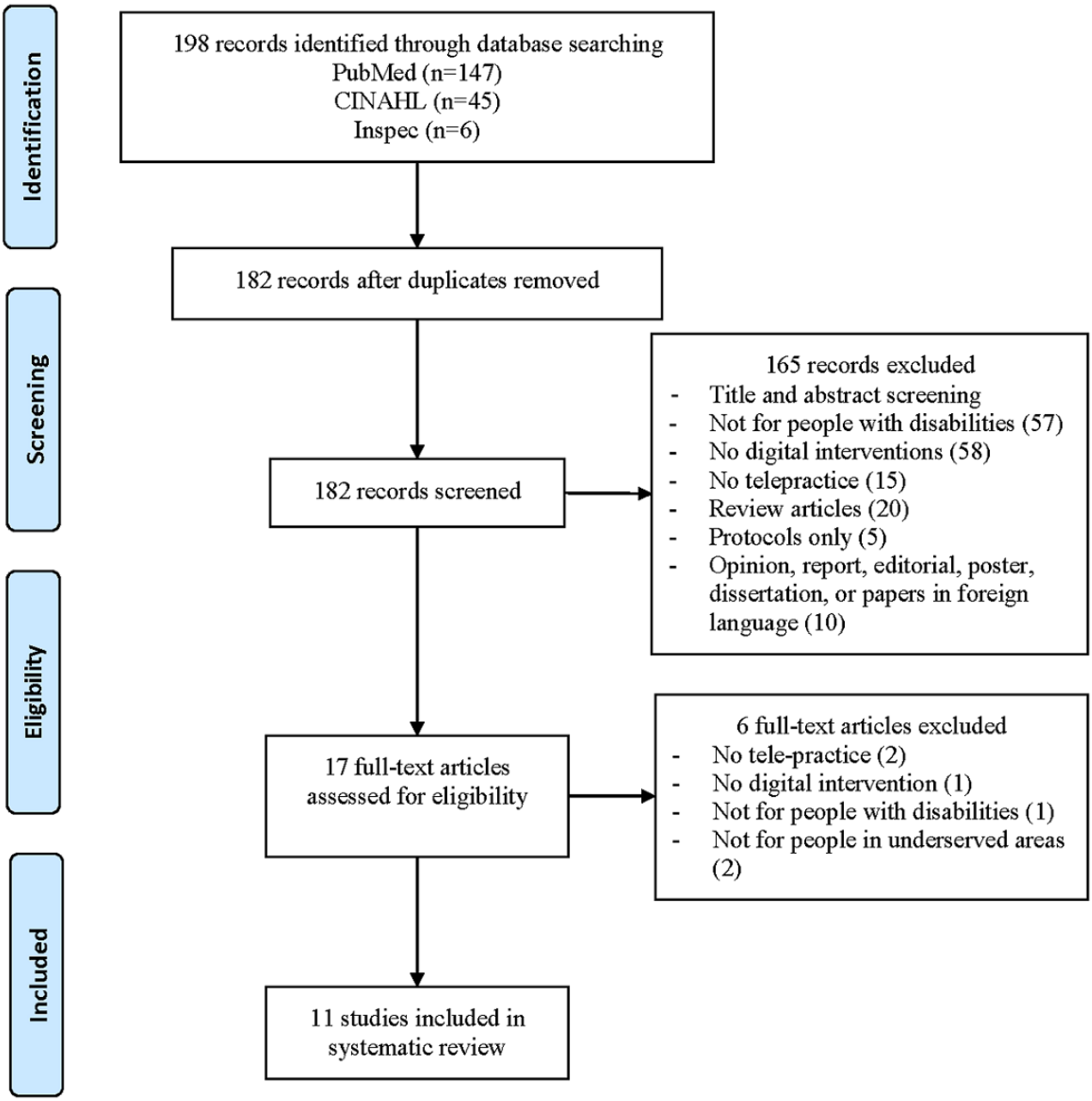
and race; study design; duration of intervention; intervention; telehealth technology (store-and-forward, teleconferencing, mobile phone app, or other approach); comparator (if applicable); outcome measures; study results; and location of the study (country or state).

Results

Study Selection

In the first round of the study selection, 16 duplicates were removed from the study list. In the second round, 165 articles were removed from the study list because they were an opinion paper (1), published in foreign language (2), a dissertation (2), an editorial (3), not for people with disabilities (57), did not include any digital interventions (58), did not use telehealth practice (15), a poster (1), a protocol (5), a report (1), or a review article (20). Each count here is only for a violation of one item in the selection criteria to avoid double count. In the third round, 6 articles were removed from the remaining 17 articles because they did not include telepractice (2), did not have any intervention (1), were not for people in underserved areas (2), or were not for people with disabilities (1). Therefore, at the end of the study selection, a total of 11 articles remained. A flowchart for this study selection is shown in Figure 1.

Figure 1. Flowchart of the study selection.



Quality Assessment

The quality assessment results shown in Table 1 [40-50] illustrate that these 11 studies met the quality criteria for being included in this systematic review. One common problem of these studies is that none of them justified their sample size (item 3b in Textbox 1). The other common problem in more than half of these studies (7/11, 64%) is that authors did not report their results in terms of statistical significance (item 6a).

Table 1. Quality assessment summary from the modified McMaster critical review form.

Study	Study design	Items on the modified McMaster critical review form										Score, n (%)
		1	2	3a	3b	4	5	6a	6b	6c	7	
Clark et al, 2002 [40]	Case study	Y[a]	Y	Y	N[b]	Y	Y	N	Y	Y	Y	8 (80)
Forducey et al, 2003 [41]	Case study	N	Y	Y	N	Y	Y	N	N	Y	Y	6 (60)
Barlow et al, 2009 [42]	Cohort	Y	Y	Y	N	Y	Y	Y	Y	N	Y	8 (80)
Kelso et al, 2009 [43]	Case study	Y	Y	Y	N	Y	N	N	Y	Y	Y	7 (70)
Schein et al, 2010 [44]	Cohort	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	9 (90)
Olsen et al, 2012 [45]	Cohort	Y	Y	Y	N	Y	N	Y	Y	Y	Y	8 (80)
Crotty et al, 2014 [46]	Cohort	Y	N	Y	N	Y	Y	N	Y	N	Y	6 (60)
Levy et al, 2015 [47]	Cohort	Y	Y	Y	N	Y	N	Y	Y	Y	Y	8 (80)
Langkamp et al, 2015 [48]	Case study	Y	Y	Y	N	Y	N	N	Y	N	Y	6 (60)
Sangelaji et al, 2017 [49]	Case study	Y	Y	Y	N	Y	Y	N	Y	N	Y	7 (70)
Portaro et al, 2018 [50]	Case study	Y	Y	N	N	Y	Y	N	Y	Y	Y	7 (70)

[a]Y=yes.

[b]N=no.

Study Characteristics

Journals

A total of 11 studies were included in this review [40-50]. All of them were published in peer-reviewed journals. Each of the following journals contained 1 (9%) of the studies: Journal of Neurologic Physical Therapy [40], NeuroRehabilitation [41], International Journal of Telerehabilitation [42], Infants and Young Children [43], Assistive Technology [44], Volta Review [45], Journal of Telemedicine and Telecare [46], Journal of Rehabilitation Research and Development [47], Telemedicine Journal and E-Health [48], European Journal of Physiotherapy [49], and Disability and Health Journal [50].

Study Locations

A total of 7 (64%) studies were performed in the United States (2 in Oklahoma [40,41], 1 in "a large western state" [43], and 1 in each of the following states: Pennsylvania [44], Utah [45], Florida [47], and Ohio [48]). The other 4 (36%) studies were from Australia [46], New Zealand [49], Italy [50], and Canada [42].

Location of Study Participants

A total of 10 (91%) studies were performed with study participants in rural or remote areas, and one (9%) was conducted with participants in both a rural nursing home and in an urban community [46].

Year Studies Were Published

There was 1 study published in each of following years: 2002, 2003, 2010, 2012, 2014, 2017, and 2018. There were 2 studies published in 2009 and 2015.

Disabilities or Disability-Causing Disorders

In 5 (45%) of the studies, the specific disabilities were mentioned, 3 were DD [43,45,48] and 2 were mobile impairment (MI) [42,44]. In the other 6 (55%) studies, the cause of the disabilities were mentioned instead, 2 were stroke [40,46], 2 were MS [47,49], 1 was TBI [41], and 1 was facio-scapulo-humeral muscular dystrophy (FSHD) [50]. It must be noted that one disease may cause multiple disabilities, for instance, stroke may cause cognitive, language, and mobility impairments. Similarly, one type of disability may be caused by different medical problems. For instance, both TBI and FSHD can lead to mobility impairment. This fact made it difficult to combine numbers from these 11 studies and perform quantitative analysis.

Telehealth Technologies

In 9 (82%) studies, various types of VC systems were used to deliver different types of interventions. In 1 (9%) study, both a website and a VC system were used to deliver the interventions [49]. In one other (9%) study, a store-and-forward technology was used to deliver the intervention [48].

Type of Studies

A total of 6 studies (55%) were small-scale case studies [40,41,43,48-50], and the other 5 (45%) were cohort studies. In most of these studies, a pre- and postevaluation was performed to determine the effectiveness of the telehealth-based interventions. Questionnaires, interviews, and focus groups were used to collect data from the study participants. A few studies also included objective evaluation, such as patients' functional level.

Sample Sizes

In the 6 case studies, the sample sizes were 1, 1, 4, 4, 4, and 4. Note that in one of these studies, there were 137 participants, but only 4 cases were described in detail [48]. In the 5 cohort studies, the sample sizes were 10, 26, 30, 40, and 104. In the last study, not all participants were people with disabilities [46]. The number of people with disabilities in the study was approximately 80.

Participant Characteristics

Not all of the studies provided the age information of their participants. In general, the participants' age in these studies was at one of 2 extremes, either children aged 0 to 3 years or people older than 50 years. According to the information provided in the studies, the average age of the reported adult participants was approximately 60 years. Similarly, not all of the studies provided the gender information of their participants. According to the studies that did report gender for adult participants, it seems that the overall gender distribution in these 11 studies was balanced. The gender of the children in these studies was not reported. Most of the studies also did not report the race of the participants. The ones that did indicated that most of the participants were white.

Potential Risk of Bias

There are various types of potential risk of bias, for instance, small sample size, limited population, gender bias, being geographically limited, age bias, education bias, and racial bias. Of the 11 studies, 7 (64%) had a small sample size. As mentioned in the previous paragraph, most of the indicated study participants were white, and therefore, there was racial bias in those studies. In 5 studies (5/11, 45%), participants were either mainly male or mainly female, and therefore, there was gender bias in those studies. Although all of these studies were for people with disabilities and the participants were recruited from one or a few geographical areas, the 6 of 11 studies that were case studies have higher risk of bias because of the limited population. Only 3 studies (3/11, 27%) had good sample sizes (40, 30, 104).

These study characteristics are summarized in Table 2.

Table 2. Characteristics of the 11 selected studies.

Reference	Disease or disability	Sample size	Participants' demographics	Outcome data collection method	Study location	Potential bias
Clark et al, 2002 [40]	Stroke	1	52 years, white, woman	Pre-post comparison	United States	Small sample size, racial bias, and gender bias
Forducey et al, 2003 [41]	Traumatic brain injury	1	39 years, man	Pre-post comparison	United States	Small sample size and gender bias
Barlow et al, 2009 [42]	MI[a]	10	Mean age=72.2 years, 8 women	Telehealth vs face-to-face comparison	Canada	Small sample size and gender bias
Kelso et al, 2009 [43]	DD[b]	4	Children (birth to 3 years)	Questionnaire and interview	United States	Small sample size
Schein et al, 2010 [44]	MI	40	Mean age=55 years, 36 white, 25 women	Pre-post comparison	United States	Racial bias and good sample size
Olsen et al, 2012 [45]	DD	30	Children (birth to 3 years)	Pre-post comparison	United States	Good sample size
Crotty et al, 2014 [46]	35 stroke, 10 fracture, 33 cognitive impairment, 4 joint replacement, 22 others	104	Community residents (n=61): mean age=73.4 years, 26 women. Rural nursing home patients (n=43): mean age=83 years, 30 women	Pre-post comparison, questionnaire, interview, and focus group	Australia	Good sample size
Levy et al, 2015 [47]	21 musculoskeletal disorder, 4 MS[c], 1 stroke	26	24 men, 18 aged 50-64 years, 8 aged >64 years	Pre-post comparison	United States	Gender bias and age bias
Langkamp et al, 2015 [48]	DD	4/137	Mean age=9.2 years, 131 white	Pre-post comparison, survey, and interview	United States	Small sample size, age bias, and racial bias
Sangelaji et al, 2017 [49]	MS	4	Ages 56, 56, 65, 75 years, all women	Pre-post comparison, interview, and questionnaire	New Zealand	Age bias and gender bias
Portaro et al, 2018 [50]	Facio-scapulo-humeral muscular dystrophy	4	4 siblings, no gender or age information	Pre-post comparison	Italy	Small sample size and limited population

[a]MI: mobile impairment.

[b]DD: developmental disability.

[c]MS: multiple sclerosis.

Results of Individual Studies

Telehealth Interventions

EI for children with DD was described in 2 studies (18%) [43,45]; PT, OT, SLT, and psychology services were provided to people with stroke in 2 studies (18%) [40,46] and people with TBI in 1 study (9%) [41]; PT services were delivered to people with MS in 1 study (9%) [47]; telemonitoring, psychological consultation, and neurological and pneumological assessment services were given to 4 siblings with FSHD in 1 study (9%) [50]; typical primary care was provided to children with DD in 1 study (9%) [48]; and assessment and prescription of wheelchair and seating were provided to people with MI in 2 studies (18%) [42,44]. The duration of the interventions ranged from a few hours to 2 years.

Outcome Measures

In most of the studies (10, 91%), the outcome measure included participants' (patients, care givers, local clinicians, and remote clinicians) satisfaction. In some studies, the outcome measures also included one of the following items: physical function, mental status, communication skills, self-care ability, cost and time savings, service time, number of hospital admissions, and goal attainment. The outcome measures, duration, and intervention results are summarized in Table 3.

Table 3. Duration, intervention, outcome measures, and intervention results for the 11 selected studies.

Reference	Duration of intervention	Interventions	Outcome measures	Results
Clark et al, 2002 [40]	17 months	PT[a], OT[b], SLT[c], vocational rehabilitation, and psychological services	Mobility, self-care ability, emotion, language ability, and cost and travel savings	Patient was functionally independent in household walking and self-care; functional use of affected lower extremity for support and balance; patient could express basic needs independently, communicate complex ideas; caregiver's mood was more positive; and cost and travel savings
Forducey et al, 2003 [41]	24 weeks	PT, OT, SLT, neuro-psychological services, and telementoring	Physical and cognitive function of patients and nursing home staff's perception and satisfaction	Improvements in neuropsychological status and physical functioning and the telementoring program was very beneficial
Barlow et al, 2009 [42]	2 years	Wheelchair seating assessment and intervention	Patient and therapists' satisfaction, intervention goal attainment, travel expense, therapists' time spent in providing service, and wait time and completion time	Clients had similar satisfaction ratings to those seen F2F[d]; clients had their goals met as often as clients seen F2F; travel cost savings; rural therapists spent more time in preparation and follow-up; and clients had shorter wait times for assessment than rural F2F clients
Kelso et al, 2009 [43]	1 month for 2 families, 3 months for the other 2 families	EI[e] (SLT, OT, and PT)	Parental satisfaction, usability of the system, interventionists' feedback, and cost and travel savings	Videoconferencing-based tele-EI system is both usable and satisfactory to most participants; parents and therapists experienced technical problems; and cost savings for delivering EI via telehealth
Schein et al, 2010 [44]	88 min on average	Assessment and prescription of wheelchair and seating	Users' satisfaction, comfort and time and cost savings	A high level of patient satisfaction and saved money and time
Olsen et al, 2012 [45]	1 year	EI, home visits, and coaching model	Cost savings, participants' rating, and provider and family satisfaction	Cost savings and increased availability of services from specialists; parents' comfort with technical skills was high; provider's ratings of comfort with the telehealth experience were high; parents were satisfied with each visit modality; most providers (79%) were satisfied with the telehealth experience; and telehealth removed time and travel barriers and increased availability of qualified personnel
Crotty et al, 2014 [46]	Up to 8 weeks	Coaching model, feedback and homework for the patient, SLT, OT, PT, and medical reviews	Participants' satisfaction, goal attainment, number of home visits, service time, travel time; cognitive impairment, mood, quality of life, and functional level and perceived ease of technology use	Participants achieved 75% of the goals; high levels of satisfaction; a 50% reduction in home visits by staff; speech therapists doubled occasions of services and direct patient contact time but halved their travel time; patients achieved >50% of their goals; most patients achieved their anticipated or better outcome; telehealth was acceptable and perceived positively by older people; and in approximately 2/3 cases, clinicians were equally satisfied with telehealth compared with F2F sessions
Levy et al, 2015 [47]	On average 99 days	PT	Functional level, quality of life, and satisfaction	Significant improvement in most outcome measures; 96% of patients were satisfied with the telehealth experience; and avoided travel miles, driving time, and travel reimbursement
Langkamp et al, 2015 [48]	1 year	Connection to primary doctor	Parents' satisfaction, school staff's satisfaction and comfort with the program, and participating practice members' experience with the program	Most parents had a high level of satisfaction with the program; parents were satisfied with the care their child received; school staff noticed benefits of telehealth; and participating providers agreed to continue the participation
Sangelaji et al, 2017 [49]	24 weeks	12 weeks Web-based physiotherapy followed by 12 weeks behavioral change intervention	Participants' feedback, physical activity, body function and composition; quality of life, fatigue, and mental status	Intervention was not effective for the participants; accepted telehealth practice; overall dissatisfaction with using the activity monitors; and both positive and negative aspects of website use

Reference	Duration of intervention	Interventions	Outcome measures	Results
Portaro et al, 2018 [50]	6 months	Telemonitoring, psychological consultation, neurological, and pneumological assessment	Number of hospital admissions, patients' satisfaction, the clinical impact, and quality of life	Reduced hospital admissions; patients had a mild improvement in emotional and mood status; body mass index remained stable; patients developed better skills to solve problems; no change on caregiver burden; and reasonable level of satisfaction

[a]PT: physical therapy.

[b]OT: occupational therapy.

[c]SLT: speech and language therapy.

[d]F2F: face-to-face.

[e]EI: early intervention.

Discussion

Principal Findings

Long-term and highly skilled therapists in various fields (such as PT, OT, SLT, and psychotherapy) are not available in underserved areas, including rural areas, remote areas, some poor urban communities, and developing countries. Telehealth may be a viable approach for delivering intervention digitally to people with disabilities in such underserved areas.

This systematic review showed that most patients had a positive opinion regarding digital intervention delivered via telehealth. Most of them had reasonable levels of satisfaction; some of them had functional improvement in motor performance, language ability, and self-care skills. Their mental status and quality of life showed improvement in some studies. In addition, telehealth made it possible for them to access desired interventions and saved them time and money.

A few studies included in this review provided services and evaluated the situations of family caregivers [40,50]. The results indicated that these caregivers were helped by the digital intervention (such as psychotherapy and communication skills) and that caregivers were satisfied with the intervention delivered to patients via telehealth.

Some studies also assessed the local and remote care providers' experience with participating in the telehealth-based intervention [46,48]. Overall, these care providers were generally satisfied with this digital intervention delivery approach as it provided intervention results comparable with face-to-face visits, increased patient contact time, and reduced travel time and costs.

The majority (6, 54%) of the 11 studies were small-scale case studies, and the rest were relatively larger-scale cohort studies. None of them were randomized controlled trials. Most of these studies used pre-post evaluation, questionnaire, and interview to determine the outcome of the intervention. They did not offer comparison with the outcomes of traditional face-to-face intervention.

For some specific interventions, such as EI, it is known that for children it is beneficial to be delivered within the child's natural environment and to use daily activities with familiar people. In this case, digital intervention via telehealth might be the only plausible approach for delivering EI to children in underserved areas at a specific time and frequency.

In most of these studies, the telehealth technology was VC for synchronized intervention, in which all parties (patients, caregivers, local care providers, and the remote care team) could interact in real time. This is desired in most cases. In some circumstances, asynchronous telehealth may be superior to synchronized communication or traditional in-office visits [48], as children with DD may not cooperate when a doctor is observing. In a store-and-forward mode, children may not have the stress, and they are more likely to cooperate when having a medical exam done by a school staff they know.

Comparison With Previous Studies

There have been some other systematic reviews of telehealth or telerehabilitation in general, but these typically only focused on a specific disease, a specific age group, a specific type of outcome, or a specific geographical area [34,51]. Our systematic review covered studies performed all over the world and with all types of disabilities and disability-causes diseases. Our study included both synchronous and asynchronous interventions with all ages. The results of our systematic review and narrative analysis are consistent with those of other reviews [34].

In 2000, a systematic review of studies of patient satisfaction with telehealth reported findings in 32 studies conducted worldwide and published between 1966 and 1998 [52]. It pointed out that although all studies reported a good level of patient satisfaction, qualitative analyses determined methodological problems with all the published work, such as low sample sizes and problematic study design, which limited the generalizability of the findings in those studies [52].

In 2003, a keyword search ("telehealth, telemedicine, or telerehabilitation") in the literature returned mostly pilot studies, case studies, and feasibility studies [53]. The situation was not significantly improved in 2007 as many of the identified studies still had limitations in study design, small sample size, and no comparison with face-to-face intervention [54].

Our systematic review covered studies published between 2003 and 2018. Comparing our results with those of the previous studies, we can see that the research studies in telehealth have not significantly improved in the past two decades. The studies reviewed in this project still suffered from the same issues: small sample size and lack of comparison with face-to-face intervention.

Some people believe that the small sample size in telehealth studies is related to the availability of technology or familiarity

with the ICT used in telehealth [55,56]. The wide adoption of the internet and smart mobile devices in recent years indicate that the availability of technology is not an issue anymore. A few studies included in this review evaluated participants' familiarity with technology and its impact on the outcomes of the intervention [45,46]. The general conclusion was that older people were less familiar with technology [57] but that age and previous experience with ICT were not barriers to digital intervention via telehealth if technology training was provided before the intervention. However, it is possible that familiarity with technology might impact study participant recruitment [46].

Limitations

This systematic review contains several limitations. The keyword search did not use a controlled vocabulary (eg, Medical Subject Headings). The inclusion criterion may have excluded studies that describe digital telehealth intervention for people with disabilities in underserved areas but do not contain the exact keywords we used. Moreover, only studies with full text written in English were included in the sample, which excluded articles in non-English journals. In addition, only peer-reviewed studies published in scholarly journals and conference proceedings were included in this study; therefore, articles published in gray literature were excluded.

Several concerns regarding the selected studies and outcomes limited the overall findings of this study. The included studies had highly heterogeneous designs and used various methods to measure the outcomes of digital interventions. Therefore, it was not feasible to conduct meta-analysis or explore the impact of these studies as a group. In addition, some studies did not include clear descriptions of the recruitment process. The studies included were at risk of selection bias, and on the individual study level, there was a lack of information about potential confounding factors such as age, gender, and educational level, which possibly could have affected self-assessed outcome.

We acknowledge that the number of included studies is small, reflecting the current state of published literature relating to digital telehealth intervention for people with disabilities in underserved areas. This review may serve as a checkpoint for the development of more, larger-scale, and higher-quality digital telehealth intervention studies for people with disabilities in underserved areas by researchers in the future. The findings from this study itself are limited by the small number of studies that met the inclusion criteria and the small sample sizes involved in each study. Therefore, although the obtained results were mainly positive, because of the small sample size, they may be considered only proof of concept instead of solid and generalizable conclusions.

This study did not include a regular telephone-based intervention. This is one of the feasible approaches for providing intervention to people in underserved areas. However, in many cases, rehabilitation intervention requires demonstration of proper procedure, and that is very difficult if not impossible via a regular telephone conversation. Considering that many people in the underserved areas, including people in developing countries, have mobile phones instead of land phones and the rate of ownership is still increasing, it is believed that many people in underserved areas can have VC for telehealth-based interventions, and therefore, we believe this decision would not have led to missing any important studies.

There are several ongoing studies in this field as well as study protocols with larger sample size [58-60]. These may generate better and more convincing results in the near future.

Acknowledgments

This research was funded in part by grants from the National Institutes of Health (NIH), 1 R43 DC04246-01 and 1 R43 DC04472-01, and the National Institute on Disability, Independent Living, and Rehabilitation Research (NIDILRR), 90RE5018, 90DP0064, and 90DPGE0002. The contents of this study do not represent the views of the NIH, NIDILRR, or the United States Government.

Conflicts of Interest

None declared.

References

1. Benjamin EJ, Virani SS, Callaway CW, Chamberlain AM, Chang AR, Cheng S, American Heart Association Council on Epidemiology and Prevention Statistics Committee and Stroke Statistics Subcommittee. Heart disease and stroke statistics-2018 update: a report from the American Heart Association. Circulation 2018 Mar 20;137(12):e67-492. [doi: 10.1161/CIR.0000000000000558] [Medline: 29386200]
2. Heron M. Deaths: leading causes for 2016. Natl Vital Stat Rep 2018 Jul;67(6):1-77 [FREE Full text] [doi: 10.4135/9781412952484.n432] [Medline: 30248017]
3. Ovbiagele B, Goldstein LB, Higashida RT, Howard VJ, Johnston SC, Khavjou OA, American Heart Association Advocacy Coordinating Committee and Stroke Council. Forecasting the future of stroke in the United States: a policy statement from the American Heart Association and American Stroke Association. Stroke 2013 Aug;44(8):2361-2375. [doi: 10.1161/STR.0b013e31829734f2] [Medline: 23697546]
4. Yagura H, Miyai I, Seike Y, Suzuki T, Yanagihara T. Benefit of inpatient multidisciplinary rehabilitation up to 1 year after stroke. Arch Phys Med Rehabil 2003 Nov;84(11):1687-1691. [doi: 10.1053/s0003-9993(03)00286-7] [Medline: 14639571]

5. Rabadi MH, Rabadi FM, Edelstein L, Peterson M. Cognitively impaired stroke patients do benefit from admission to an acute rehabilitation unit. Arch Phys Med Rehabil 2008 Mar;89(3):441-448. [doi: 10.1016/j.apmr.2007.11.014] [Medline: 18295621]
6. Flansbjer UB, Lexell J, Brogårdh C. Long-term benefits of progressive resistance training in chronic stroke: a 4-year follow-up. J Rehabil Med 2012 Mar;44(3):218-221 [FREE Full text] [doi: 10.2340/16501977-0936] [Medline: 22366878]
7. Centers for Disease Control and Prevention. 2015. Traumatic Brain Injury & Concussion: Report to Congress on Traumatic Brain Injury Epidemiology and Rehabilitation URL: https://www.cdc.gov/traumaticbraininjury/pubs/congress_epi_rehab.html [accessed 2019-09-03]
8. Harvey LA, Close JC. Traumatic brain injury in older adults: characteristics, causes and consequences. Injury 2012 Nov;43(11):1821-1826. [doi: 10.1016/j.injury.2012.07.188] [Medline: 22884759]
9. Kurtzke JF. Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). Neurology 1983 Nov;33(11):1444-1452. [doi: 10.1212/wnl.33.11.1444] [Medline: 6685237]
10. Weinshenker BG, Bass B, Rice GP, Noseworthy J, Carriere W, Baskerville J, et al. The natural history of multiple sclerosis: a geographically based study. I. Clinical course and disability. Brain 1989 Feb;112(Pt 1):133-146. [doi: 10.1093/brain/112.1.133] [Medline: 2917275]
11. Lublin FD, Reingold SC. Defining the clinical course of multiple sclerosis: results of an international survey. National multiple sclerosis society (USA) advisory committee on clinical trials of new agents in multiple sclerosis. Neurology 1996 Apr;46(4):907-911. [doi: 10.1212/wnl.46.4.907] [Medline: 8780061]
12. Soldán MM, Novotna M, Zeid N, Kale N, Tutuncu M, Crusan DJ, et al. Relapses and disability accumulation in progressive multiple sclerosis. Neurology 2015 Jan 6;84(1):81-88 [FREE Full text] [doi: 10.1212/WNL.0000000000001094] [Medline: 25398229]
13. Solari A, Filippini G, Gasco P, Colla L, Salmaggi A, la Mantia L, et al. Physical rehabilitation has a positive effect on disability in multiple sclerosis patients. Neurology 1999 Jan 1;52(1):57-62. [doi: 10.1212/wnl.52.1.57] [Medline: 9921849]
14. Centers for Disease Control and Prevention. 2018. Facts About Developmental Disabilities URL: https://www.cdc.gov/ncbddd/developmentaldisabilities/facts.html [accessed 2018-10-10]
15. Boyle CA, Boulet S, Schieve LA, Cohen RA, Blumberg SJ, Yeargin-Allsopp M, et al. Trends in the prevalence of developmental disabilities in US children, 1997-2008. Pediatrics 2011 Jun;127(6):1034-1042. [doi: 10.1542/peds.2010-2989] [Medline: 21606152]
16. Scheideman-Miller C, Clark P, Hodge B, Prouty D. Rural Post-Acute Stroke Care Using Multidisciplinary Telerehabilitation. In: Proceedings of the 34th Annual Hawaii International Conference on System Sciences. 2001 Presented at: HICSS'01; January 6, 2001; Maui, HI, USA. [doi: 10.1109/hicss.2001.926563]
17. US Census Bureau. 2016. New Census Data Show Differences Between Urban and Rural Populations URL: https://www.census.gov/newsroom/press-releases/2016/cb16-210.html [accessed 2018-10-10]
18. Jones CA, Parker TS, Ahearn M, Mishra AK, Variyam JN. US Department of Agriculture - Economic Research Service. 2009. ERS Report Summary: Health Status and Health Care Access of Farm and Rural Populations URL: https://www.ers.usda.gov/webdocs/publications/44424/9370_eib57_reportsummary_1_.pdf [accessed 2018-10-10]
19. Rural Health Information Hub. 2014. Distribution of US Health Care Providers Residing in Rural and Urban Areas URL: https://www.ruralhealthinfo.org/assets/1275-5131/rural-urban-workforce-distribution-nchwa-2014.pdf [accessed 2018-10-10]
20. Iezzoni LI, Killeen MB, O'Day BL. Rural residents with disabilities confront substantial barriers to obtaining primary care. Health Serv Res 2006 Aug;41(4 Pt 1):1258-1275 [FREE Full text] [doi: 10.1111/j.1475-6773.2006.00534.x] [Medline: 16899006]
21. Davidsson N, Södergård B. Access to healthcare among people with physical disabilities in rural Louisiana. Soc Work Public Health 2016;31(3):188-195. [doi: 10.1080/19371918.2015.1099496] [Medline: 26983377]
22. MacKinney AC, Coburn AF, Lundblad JP, McBride TD, Mueller KJ, Watson SD. Rural Policy Research Institute. 2014. Access to Rural Health Care – A Literature Review and New Synthesis URL: http://www.rupri.org/wp-content/uploads/Access-to-Rural-Health-Care-A-Literature-Review-and-New-Synthesis.-RUPRI-Health-Panel.-August-2014-1.pdf [accessed 2018-10-01]
23. Ahmed SM, Lemkau JP, Nealeigh N, Mann B. Barriers to healthcare access in a non-elderly urban poor American population. Health Soc Care Community 2001 Nov;9(6):445-453. [doi: 10.1046/j.1365-2524.2001.00318.x] [Medline: 11846824]
24. Cherry CO, Chumbler NR, Richards K, Huff A, Wu D, Tilghman LM, et al. Expanding stroke telerehabilitation services to rural veterans: a qualitative study on patient experiences using the robotic stroke therapy delivery and monitoring system program. Disabil Rehabil Assist Technol 2017 Jan;12(1):21-27. [doi: 10.3109/17483107.2015.1061613] [Medline: 26135221]
25. Pew Research Center. 2019. Internet/Broadband Fact Sheet URL: http://www.pewinternet.org/fact-sheet/internet-broadband/ [accessed 2019-06-20]
26. Pew Research Center. 2019. Mobile Fact Sheet URL: http://www.pewinternet.org/fact-sheet/mobile/ [accessed 2019-06-20]
27. Finch L. Ingenious engineering. Rehab Manag 2000;13(6):104-110.
28. Schmeler MR, Schein RM, McCue M, Betz K. Telerehabilitation and clinical applications: research, opportunities, and challenges. Int J Telerehab 2015 Mar 16:12-24. [doi: 10.5195/ijt.2008.701]

29. Harper D. Telehealth. In: Roberts MC, Steele RG, editors. Handbook of Pediatric Psychology. Third Edition. New York: Guilford Press; 2003.
30. Georgeadis A, Brennan D, Barker L, Baron C. Telerehabilitation and its effect on story retelling by adults with neurogenic communication disorders. Aphasiology 2010 Jun 24;18(5-7):639-652. [doi: 10.1080/02687030444000075]
31. Hall N, Boisvert M, Steele R. Telepractice in the assessment and treatment of individuals with aphasia: a systematic review. Int J Telerehabil 2013;5(1):27-38 [FREE Full text] [doi: 10.5195/ijt.2013.6119] [Medline: 25945211]
32. Wade SL, Wolfe C, Brown TM, Pestian JP. Putting the pieces together: preliminary efficacy of a web-based family intervention for children with traumatic brain injury. J Pediatr Psychol 2005;30(5):437-442. [doi: 10.1093/jpepsy/jsi067] [Medline: 15944171]
33. Kairy D, Lehoux P, Vincent C, Visintin M. A systematic review of clinical outcomes, clinical process, healthcare utilization and costs associated with telerehabilitation. Disabil Rehabil 2009;31(6):427-447. [doi: 10.1080/09638280802062553] [Medline: 18720118]
34. Kruse CS, Krowski N, Rodriguez B, Tran L, Vela J, Brooks M. Telehealth and patient satisfaction: a systematic review and narrative analysis. BMJ Open 2017 Aug 3;7(8):e016242 [FREE Full text] [doi: 10.1136/bmjopen-2017-016242] [Medline: 28775188]
35. Laver KE, Schoene D, Crotty M, George S, Lannin NA, Sherrington C. Telerehabilitation services for stroke. Cochrane Database Syst Rev 2013 Dec 16(12):CD010255 [FREE Full text] [doi: 10.1002/14651858.CD010255.pub2] [Medline: 24338496]
36. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JP, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. Br Med J 2009 Jul 21;339:b2700 [FREE Full text] [doi: 10.1136/bmj.b2700] [Medline: 19622552]
37. Kitchenham B, Charters S. Department of Computer Science, University of Auckland. 2007. Guidelines for Performing Systematic Literature Reviews in Software Engineering URL: https://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.117.471 [accessed 2016-04-13] [WebCite Cache ID 6gk0beToT]
38. Letts L, Wilkins S, Law M, Stewart D, Bosch J, Westmoreland M. 2007. Critical Review Form - Qualitative Studies (Version 2.0) URL: https://srs-mcmaster.ca/wp-content/uploads/2015/04/Critical-Review-Form-Qualitative-Studies-Version-2-English.doc [accessed 2019-06-20]
39. Law M, Stewart D, Pollock N, Letts L, Bosch J, Westmoreland M. 1998. Critical Review Form - Quantitative Studies URL: https://srs-mcmaster.ca/wp-content/uploads/2015/04/Critical-Review-Form-Quantitative-Studies-English.pdf [accessed 2019-06-20]
40. Clark PG, Dawson SJ, Scheideman-Miller C, Post ML. Telerehab: stroke teletherapy and management using two-way interactive video. J Neurol Phys Ther 2002;26(2):87-93. [doi: 10.1097/01253086-200226020-00006]
41. Forducey PG, Ruwe WD, Dawson SJ, Scheideman-Miller C, McDonald NB, Hantla MR. Using telerehabilitation to promote TBI recovery and transfer of knowledge. NeuroRehabilitation 2003;18(2):103-111. [Medline: 12867673]
42. Barlow IG, Liu L, Sekulic A. Wheelchair seating assessment and intervention: a comparison between telerehabilitation and face-to-face service. Int J Telerehabil 2009;1(1):17-28 [FREE Full text] [doi: 10.5195/ijt.2009.868] [Medline: 25945159]
43. Kelso GL, Fiechtl BJ, Olsen ST, Rule S. The feasibility of virtual home visits to provide early intervention. Infant Young Child 2009;22(4):332-340. [doi: 10.1097/IYC.0b013e3181b9873c]
44. Schein RM, Schmeler MR, Saptono A, Brienza D. Patient satisfaction with telerehabilitation assessments for wheeled mobility and seating. Assist Technol 2010;22(4):215-222. [doi: 10.1080/10400435.2010.518579] [Medline: 21306067]
45. Olsen S, Fiechtl B, Rule S. An evaluation of virtual home visits in early intervention: feasibility of 'virtual intervention'. Volta Rev 2012 Dec;112(3):267-282. [doi: 10.17955/tvr.112.3.m.702]
46. Crotty M, Killington M, van den Berg M, Morris C, Taylor A, Carati C. Telerehabilitation for older people using off-the-shelf applications: acceptability and feasibility. J Telemed Telecare 2014 Oct;20(7):370-376. [doi: 10.1177/1357633X14552382] [Medline: 25399997]
47. Levy CE, Silverman E, Jia H, Geiss M, Omura D. Effects of physical therapy delivery via home video telerehabilitation on functional and health-related quality of life outcomes. J Rehabil Res Dev 2015;52(3):361-370 [FREE Full text] [doi: 10.1682/JRRD.2014.10.0239] [Medline: 26230650]
48. Langkamp DL, McManus MD, Blakemore SD. Telemedicine for children with developmental disabilities: a more effective clinical process than office-based care. Telemed J E Health 2015 Feb;21(2):110-114 [FREE Full text] [doi: 10.1089/tmj.2013.0379] [Medline: 25545598]
49. Sangelaji B, Smith C, Paul L, Treharne G, Hale L. Promoting physical activity engagement for people with multiple sclerosis living in rural settings: a proof-of-concept case study. Eur J Physiother 2017 Oct 14;19(sup1):17-21. [doi: 10.1080/21679169.2017.1381306]
50. Portaro S, Calabrò RS, Bramanti P, Silvestri G, Torrisi M, Conti-Nibali V, et al. Telemedicine for Facio-Scapulo-Humeral muscular dystrophy: a multidisciplinary approach to improve quality of life and reduce hospitalization rate? Disabil Health J 2018 Apr;11(2):306-309. [doi: 10.1016/j.dhjo.2017.09.003] [Medline: 28967584]
51. Tchero H, Teguo MT, Lannuzel A, Rusch E. Telerehabilitation for stroke survivors: systematic review and meta-analysis. J Med Internet Res 2018 Oct 26;20(10):e10867 [FREE Full text] [doi: 10.2196/10867] [Medline: 30368437]

52. Mair F, Whitten P. Systematic review of studies of patient satisfaction with telemedicine. Br Med J 2000 Jun 3;320(7248):1517-1520 [FREE Full text] [doi: 10.1136/bmj.320.7248.1517] [Medline: 10834899]
53. Aoki N, Dunn K, Johnson-Throop KA, Turley JP. Outcomes and methods in telemedicine evaluation. Telemed J E Health 2003;9(4):393-401. [doi: 10.1089/153056203772744734] [Medline: 14980098]
54. Whitten P, Johannessen LK, Soerensen T, Gammon D, Mackert M. A systematic review of research methodology in telemedicine studies. J Telemed Telecare 2007;13(5):230-235. [doi: 10.1258/135763307781458976] [Medline: 17697509]
55. Grigsby J, Bennett RE. Alternatives to randomized controlled trials in telemedicine. J Telemed Telecare 2006;12(Suppl 2):S77-S84. [doi: 10.1258/135763306778393162] [Medline: 16989679]
56. Scott RE, McCarthy FG, Jennett PA, Perverseff T, Lorenzetti D, Saeed A, et al. Telehealth outcomes: a synthesis of the literature and recommendations for outcome indicators. J Telemed Telecare 2007 Dec;13(Suppl 2):1-38. [doi: 10.1258/135763307782213552] [Medline: 18005523]
57. Neter E, Brainin E. eHealth literacy: extending the digital divide to the realm of health information. J Med Internet Res 2012 Jan 27;14(1):e19 [FREE Full text] [doi: 10.2196/jmir.1619] [Medline: 22357448]
58. Rimmer JH, Thirumalai M, Young H, Pekmezi D, Tracy T, Riser E, et al. Rationale and design of the tele-exercise and multiple sclerosis (TEAMS) study: a comparative effectiveness trial between a clinic- and home-based telerehabilitation intervention for adults with multiple sclerosis (MS) living in the deep south. Contemp Clin Trials 2018 Aug;71:186-193. [doi: 10.1016/j.cct.2018.05.016] [Medline: 29859267]
59. Gauthier LV, Kane C, Borstad A, Strahl N, Uswatte G, Taub E, et al. Video game rehabilitation for outpatient stroke (VIGoROUS): protocol for a multi-center comparative effectiveness trial of in-home gamified constraint-induced movement therapy for rehabilitation of chronic upper extremity hemiparesis. BMC Neurol 2017 Jun 8;17(1):109 [FREE Full text] [doi: 10.1186/s12883-017-0888-0] [Medline: 28595611]
60. Sureshkumar K, Murthy GV, Kinra S, Goenka S, Kuper H. Development and evaluation of a smartphone-enabled, caregiver-supported educational intervention for management of physical disabilities following stroke in India: protocol for a formative research study. BMJ Innov 2015 Jul;1(3):117-126 [FREE Full text] [doi: 10.1136/bmjinnov-2015-000042] [Medline: 26751379]

Abbreviations

CP: cerebral palsy
DD: developmental disability
EI: early intervention
FSHD: facio-scapulo-humeral muscular dystrophy
ICT: information and communication technology
IVR: interactive voice response
MI: mobile impairment
MS: multiple sclerosis
OT: occupational therapy
PT: physical therapy
SLT: speech language therapy
TBI: traumatic brain injury
VC: videoconferencing

Edited by G Eysenbach; submitted 28.11.18; peer-reviewed by H Tchero, M Oroojeni Mohammad Javad, M Javad, C Kruse; comments to author 31.03.19; revised version received 07.08.19; accepted 21.08.19; published 25.10.19.

Please cite as:
Zhou L, Parmanto B
Reaching People With Disabilities in Underserved Areas Through Digital Interventions: Systematic Review
J Med Internet Res 2019;21(10):e12981
URL: http://www.jmir.org/2019/10/e12981/
doi:10.2196/12981
PMID:31654569

Original Paper

Participant-Centered Online Active Surveillance for Adverse Events Following Vaccination in a Large Clinical Trial: Feasibility and Usability Study

Sally-Anne Munnoch[1*], MAppEpi; Patrick Cashman[1,2*], MPH&TM; Roseanne Peel[2*], MPH; John Attia[1,2*], PhD; Alexis Hure[2,3*], PhD; David N Durrheim[1,2,3*], FAFPHM

[1]Hunter New England Local Health District, Wallsend, Australia

[2]University of Newcastle, Newcastle, Australia

[3]Hunter Medical Research Institute, New Lambton Heights, Australia

[*]all authors contributed equally

Corresponding Author:
Sally-Anne Munnoch, MAppEpi
Hunter New England Local Health District
Locked Mail Bag 10
Wallsend, NSW, 2287
Australia
Phone: 61 429231908
Email: sally.munnoch@health.nsw.gov.au

Abstract

Background: Active participant monitoring of adverse events following immunization (AEFI) is a recent development to improve the speed and transparency of vaccine safety postmarketing. Vaxtracker, an online tool used to monitor vaccine safety, has successfully demonstrated its usefulness in postmarketing surveillance of newly introduced childhood vaccines. However, its use in older participants, or for monitoring patients participating in large clinical trials, has not been evaluated.

Objective: The objective of this study was to monitor AEFIs in older participants enrolled in the Australian Study for the Prevention through the Immunisation of Cardiovascular Events (AUSPICE) trial, and to evaluate the usefulness and effectiveness of Vaxtracker in this research setting.

Methods: AUSPICE is a multicenter, randomized, placebo-controlled, double-blinded trial in which participants aged 55 to 61 years were given either the pneumococcal polysaccharide vaccine (23vPPV) or 0.9% saline placebo. Vaxtracker was used to monitor AEFIs in participants in either treatment arm through the administration of two online questionnaires. A link to each questionnaire was sent to participants via email or short message service (SMS) text message 7 and 28 days following vaccination. Data were collated and analyzed in near-real time to identify any possible safety signals indicating problems with the vaccine or placebo.

Results: All 4725 AUSPICE participants were enrolled in Vaxtracker. Participant response rates for the first and final survey were 96.47% (n=4558) and 96.65% (n=4525), respectively. The online survey was completed by 90.23% (4083/4525) of Vaxtracker participants within 3 days of receiving the link. AEFIs were reported by 34.40% (805/2340) of 23vPPV recipients and 10.29% (240/2332) of placebo recipients in the 7 days following vaccination. Dominant symptoms for vaccine and placebo recipients were pain at the injection site (587/2340, 25.09%) and fatigue (103/2332, 4.42%), respectively. Females were more likely to report symptoms following vaccination with 23vPPV compared with males (433/1138, 38.05% versus 372/1202, 30.95%; P<.001).

Conclusions: Vaxtracker is an effective tool for monitoring AEFIs in the 55 to 61 years age group. Participant response rates were high for both surveys, in both treatment arms and for each method of sending the survey. This study indicates that administration of 23vPPV was well-tolerated in this cohort. Vaxtracker has successfully demonstrated its application in the monitoring of adverse events in near-real time following vaccination in people participating in a national clinical trial.

Trial Registration: Australian New Zealand Trial Registry Number (ACTRN) 12615000536561; https://www.anzctr.org.au/Trial/Registration/TrialReview.aspx?id=368506

(J Med Internet Res 2019;21(10):e14791) doi:10.2196/14791

KEYWORDS

clinical trials; active surveillance; adverse events following immunization; technology; vaccination

Introduction

Active participant-centered monitoring of adverse events following immunization (AEFI) is a recent development to improve the speed and transparency of vaccine safety postmarketing [1]. An online vaccine safety monitoring tool called Vaxtracker has demonstrated its effectiveness in active postmarketing surveillance of AEFIs in children vaccinated with the seasonal influenza vaccine [2,3]. Parents or carers of recently vaccinated children who agreed to participate were sent an email and/or SMS (short message service) text messaging with an embedded hyperlink. On responding to the link, the recipient of the message was directed to a simple, individualized online questionnaire, which could be completed via mobile phone or computer. The time between vaccination and message receipt was based on the vaccine type (live versus inactivated vaccines) and known onset time of possible symptoms associated with the vaccine. Completion of the online survey allowed for timely collation, analysis, and reporting of AEFI data, as well as detecting signals indicating possible safety concerns associated with a vaccine [2].

Vaxtracker was used to monitor AEFIs in study participants enrolled in the Australian Study for the Prevention through Immunisation of Cardiovascular Events (AUSPICE). The primary objective of AUSPICE is to determine whether the pneumococcal polysaccharide vaccine (23vPPV) is protective against cardiovascular events (fatal and nonfatal acute coronary syndrome and ischemic strokes) [4]. Participants aged 55 to 60 years at the time of recruitment with no history of cardiac or stroke events but who reported risk factors for a future cardiovascular event (ie, obesity, hypertension, or hypercholesterolemia) were recruited. Participants were randomly allocated in a double-blind trial design to either the active (23vPPV) or control (0.9% saline) treatment arms. AUSPICE participants are being followed over 6 years to determine the incidence of cardiovascular events and to compare antibody titers in response to the vaccine.

Vaxtracker has demonstrated success in safety surveillance of newly introduced childhood vaccines, including utilization and acceptance by the parents or carers of young children [2,3,5], but it has not been widely used by older populations. AUSPICE provided the opportunity to test Vaxtracker's functionality in an older cohort of Australians, aged 55 to 60 years, participating in a large clinical trial. In Australia, 23vPPV is routinely administered to non-Indigenous adults, who are not at an increased risk of invasive pneumococcal disease, at age 65 years. Generally, 23vPPV is well-tolerated among recipients in this age group receiving the vaccine for the first time, although AEFI incidence and severity are known to increase with additional doses of the vaccine [6-8]. Symptoms commonly reported following vaccination include injection site reactions (pain, redness, swelling) and systemic events (fatigue, headache, low fever) [6,8,9]. Serious AEFIs, including cellulitis and swelling from joint to joint, have also been reported with the first and subsequent vaccine doses [7-9]. The age group for participation in AUSPICE was selected to ensure that an older cohort did not receive placebo instead of age-appropriate vaccine [4] and the risk factors required for inclusion in the study did not overlap with current recommendations for vaccine administration to prevent invasive pneumococcal disease. The vaccine is not routinely given to the 55 to 60 year age group; therefore, postlicensure safety assessments for this age group are limited, with a single study reporting injection site reactions being more common in a younger cohort (50-64 year age group) than in the usually targeted age group those aged 65 years and older [6]. Thus, monitoring the participants for adverse events following vaccination with either 23vPPV or placebo was crucial to ensure the safety of participants in AUSPICE.

There are two aims to our study: (1) to identify adverse events following vaccination with either 23vPPV or saline placebo in older persons participating in a large national clinical trial, and (2) to evaluate the usefulness and effectiveness of Vaxtracker in this research setting.

Methods

Study Design

The study design for AUSPICE was a multicenter, randomized, placebo-controlled, and double-blinded clinical trial. Randomization was stratified by sex and center in a 1:1 ratio for 23vPPV and 0.9% saline placebo [4]. Letters inviting people to participate in AUSPICE were sent to a random selection of residents aged 55 to 60 years who resided within a 25 km radius of one of six study sites in Australia by Medicare, Australia's national health insurance provider [4]. To manage recruitment, enrollment, and vaccination of eligible participants, the mail-out of letters inviting people to participate was staggered over the first 22 months of the study period, commencing February 2016.

Interested participants completed an online or paper-based screening questionnaire based on study inclusion and exclusion criteria. On receipt of the completed questionnaire by the research team, participants were invited to attend their closest study clinic where their demographics, self-reported medical history, and medication information were collected, and study eligibility verified. Potential study participants were asked whether they had a previous history of pneumococcal vaccination; if so, they were excluded from the study [4].

On confirmation of eligibility, participants were allocated to one of two treatment arms via a Web-based randomization system. The active vaccine or placebo was administered intramuscularly in the deltoid region by a registered immunization nurse. Syringes for both the vaccine and placebo were similar in appearance to ensure blinding of participants [4].

Vaxtracker

Vaxtracker was used to monitor AUSPICE participants for the occurrence of possible adverse events following vaccination with either 23vPPV or 0.9% saline solution. At the time of

vaccine or placebo administration, participants were provided with information on the Vaxtracker component of AUSPICE and were encouraged to report the occurrence of immediate severe adverse events to the study site staff and to complete the Vaxtracker surveys once received. Staff emphasized the importance of completing the survey even when no symptoms were experienced following vaccination. Following double-blinded allocation and administration of the vaccine or placebo, basic demographic information, contact telephone number and email address, and vaccination details of study participants were transferred automatically from the AUSPICE database into Vaxtracker.

Seven days after vaccination, all participants enrolled in Vaxtracker were sent an email and/or SMS text message containing an embedded link that directed them to an individualized Vaxtracker survey (first survey). Seven days was used to identify known serious but rare AEFIs following vaccination with 23vPPV, including severe cellulitis and anaphylaxis, in a timely manner without compromising recall by participants [8]. This online survey requested participants to confirm their demographic details, note any chronic medical conditions, other vaccines received in the 7 days before or following the study vaccination, and to note if any symptoms were experienced in the 7 days following vaccination. If a person noted "yes" to this question, they were asked whether they had experienced any of the following 13 symptoms: redness at injection site, moderate swelling at the injections site, major swelling at the injection site (elbow joint to shoulder joint), pain at injection site, limitation of arm movement, fever, fatigue, headache, chills, rash, generalized muscle pain, generalized joint pain, or other symptoms (free text response). Noting "yes" to a symptom prompted additional symptom-specific questions, including size of swelling or redness, severity assessment of pain, extent of arm movement limitation, fever temperature, or whether the joint or muscle pain was new or aggravated. Participants were also asked whether they sought medical advice regarding the symptom selected.

Vaxtracker participants who had responded to the first survey were sent a final survey either 28 days following vaccination or immediately after submission of the first survey if the first survey was completed more than 28 days following vaccination. The final survey, again sent by either email and/or SMS text message, was designed to identify severe health events in the month following vaccination with either the active or placebo vaccine. This questionnaire asked participants whether they had been hospitalized in the previous 4 weeks and, if so, requested the admission date, diagnosis, and a preferred contact number so they could be interviewed further by study staff. All participants responding to the final survey were asked to note any other comments. We again used 28 days to identify severe but rare adverse events following vaccination with 23vPPV [6,8,10].

Participants who had not completed a survey (nonresponders) were sent two messages 3 and 6 days after the initial survey dispatch through Vaxtracker's automated reminder program. If no response was received within 3 days of the second reminder, study site staff attempted to contact the nonrespondents by telephone.

It was recognized that a small proportion of the study population might not have access to a mobile phone or email address. Telephone interviews were conducted by study staff for Vaxtracker participants providing a landline phone number, using the same questionnaire. Interviewers were blinded to the vaccine status of the participants.

Questionnaire Design

Symptoms used for the Vaxtracker questionnaire and time periods for message dispatch were identified through a literature review of possible adverse events following vaccination with 23vPPV [6,9-11].

The Vaxtracker surveys were designed to ensure easy navigation using a computer mouse, stylus, or finger, to cater for people using either a desktop or laptop computer, tablet, or mobile phone (Figures 1 and 2). A combination of drop-down fields, radio buttons, calendar control fields, and minimal free text fields were used to improve data quality.

Figure 1. Screenshot of the main Vaxtracker survey accessed from an email via a desktop computer.

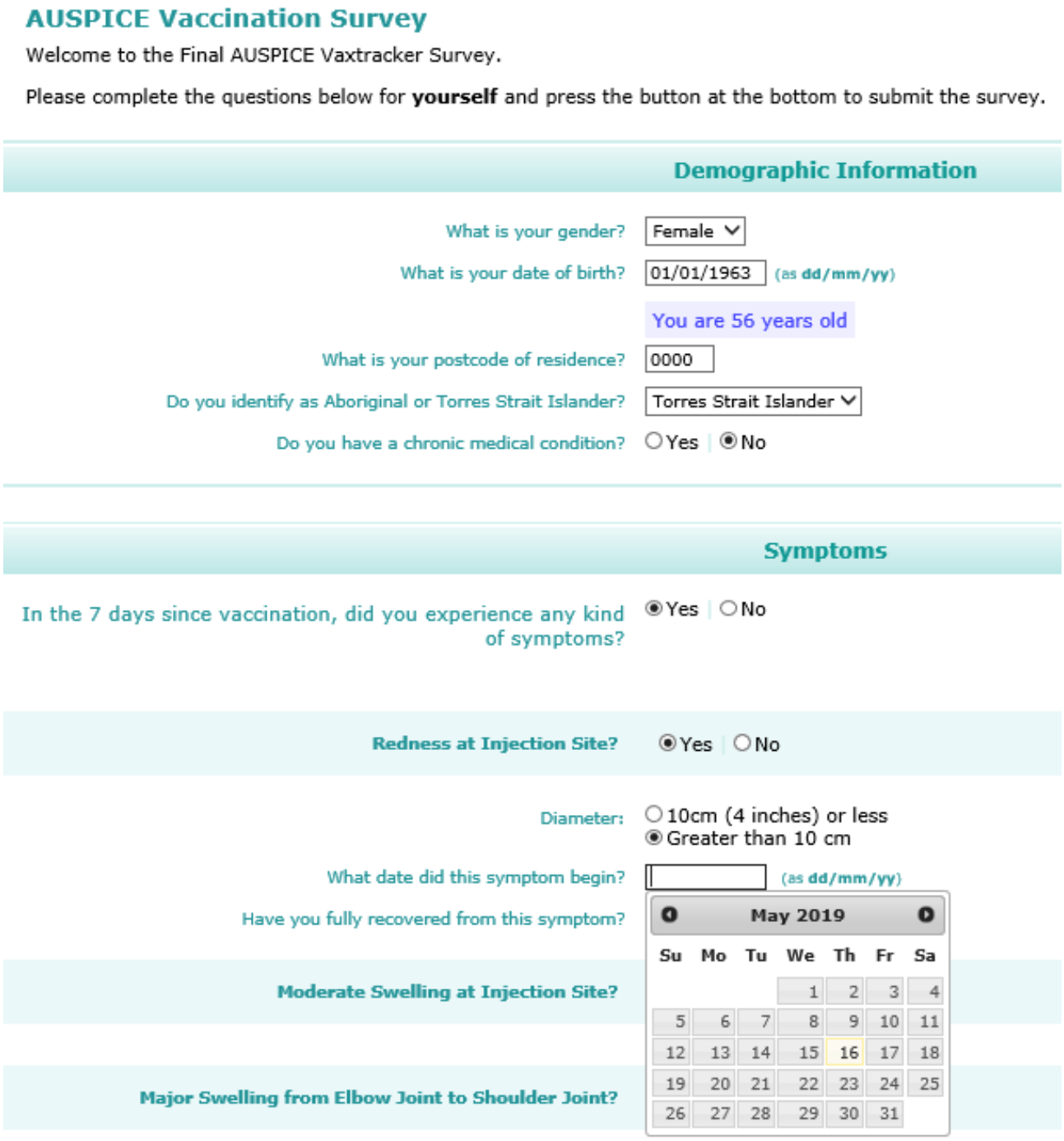
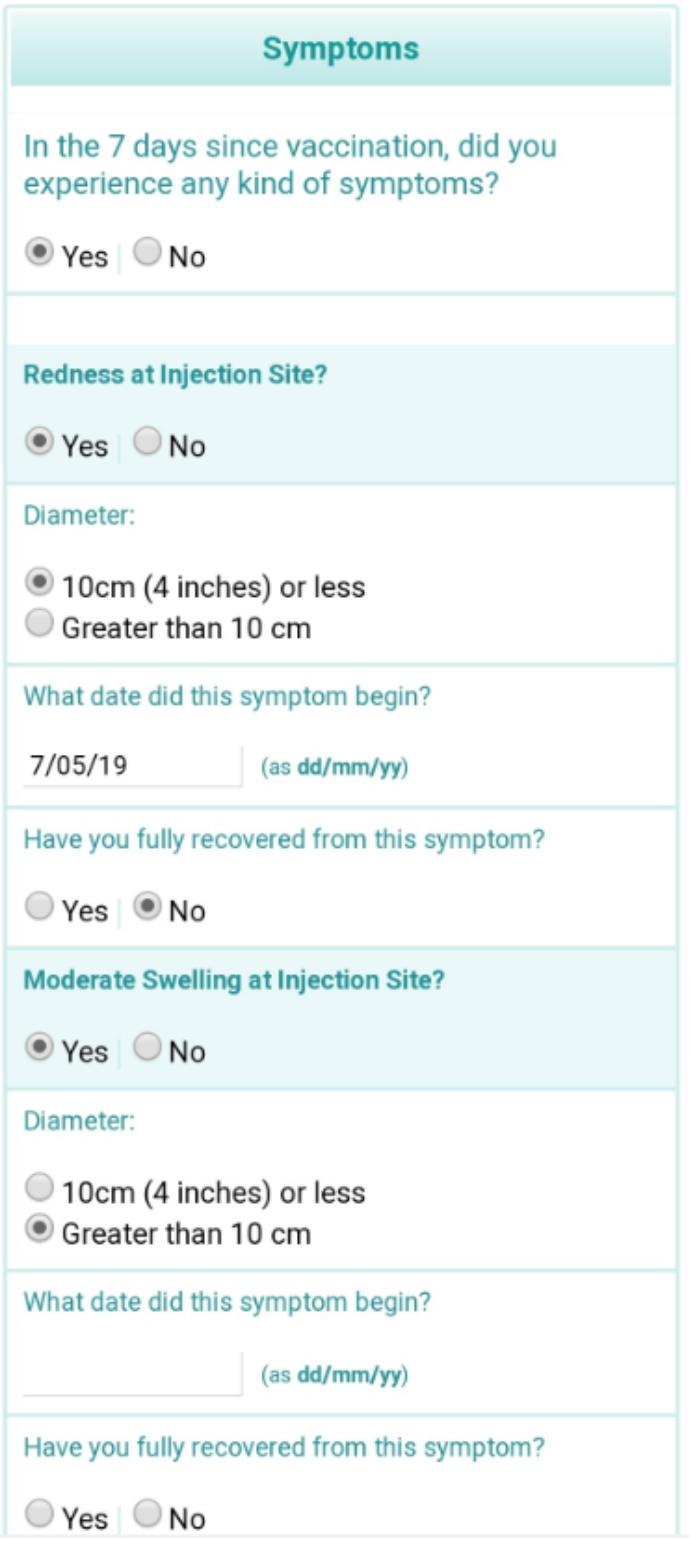


Figure 2. Screenshot of main Vaxtracker Survey viewed from a mobile phone.



Severe Adverse Event Reporting and Follow-Up

AUSPICE brochures, the study's nurse immunizers, and the Vaxtracker emails, SMS text messages, and questionnaires encouraged participants concerned about the severity of their symptoms following vaccination to seek medical advice immediately. In this instance, participants were requested to provide information about their participation in AUSPICE and possible vaccination with 23vPPV.

When a patient reported a serious symptom or medical event (extensive limb swelling, visit to an emergency department or hospital, or other serious events) via the online questionnaire, an automated serious symptom alert was forwarded to the Vaxtracker team and study staff at the center where the patient had received their vaccination. These participants were then contacted by telephone by study staff and interviewed to confirm the serious symptom, the outcome of the visit, and to provide additional counseling and information as required.

Signal Detection

To identify possible safety signals associated with 23vPPV used for AUSPICE, we compared individual reaction rates with data reported by Jackson et al [10] comparing the safety of the 13-valent pneumococcal conjugate vaccine to the 23vPPV in pneumococcal vaccine-naïve adults. Although the age of our cohort was younger (55-60 years compared with 60-64 years) and the time period for symptom review was shorter (7 days compared with 14 days), these data were considered useful for the purposes of signal identification.

Incorrect Contact Detail Monitoring

Email and SMS text message dispatch records were monitored weekly to identify incorrect email addresses. A dedicated Vaxtracker mailbox was used to receive "bounced" emails. Recipients of the messages were asked to respond to the Vaxtracker site if the message was received erroneously. If an error in an email address or mobile phone number was identified, survey completion logs were reviewed to check whether the survey had been completed through the alternate link, and the participant was contacted to correct the error. If the surveys had not been completed, study staff were asked to confirm contact details for participants, and if no error was identified, to contact the AUSPICE participant directly using alternative contact details (ie, home or work landline number).

Vaxtracker Platform

The AUSPICE Vaxtracker Web-based application is built on the ASP.NET MVC framework and related Microsoft platforms [12]. It used the Application Programming Interface (API) functionality of the app, allowing for secure transfer of AUSPICE participant information to the Vaxtracker database. A unique identifier enabling the linking of AUSPICE participants to Vaxtracker was generated, ensuring that duplicates would not be inadvertently created via a system or communications error. For each enrollment request, Vaxtracker returned its own unique identifier for the new participant back to AUSPICE, permitting full two-way cross-linkage between the datasets of the two systems.

Analyses

Descriptive analyses and statistical calculations were conducted in Excel 2010 (Microsoft Corporation, Redmond, WA, USA) and Stata 15 (StataCorp LP, College Station, TX, USA). Differences in proportions were compared using chi-square tests. A .05 level of significance was used for all analyses.

Ethics and Funding

The University of Newcastle (H-2014-0064) and Hunter New England Human Research Ethics Committees (15/08/19/3.01) were the lead ethics committees for the AUSPICE trial, including approval for the monitoring of adverse events using Vaxtracker. Further ethics approvals or registration occurred with each of the local HRECs for all AUSPICE study centers.

AUSPICE was funded by the National Health and Medical Research Council of Australia.

Results

We enrolled all 4725 AUSPICE participants in the Vaxtracker component of the study. The period for study enrollment was from February 2016 to December 2017. The median age of participants in Vaxtracker was 58.1 years (range 55-61 years) with males representing 51.47% (2432/4725) of the cohort. A small number (n=38) of participants older than 60 years were recruited into the study because their birthday fell within the period between study registration and vaccination. There were slightly more participants randomized to the treatment arm of the trial than the placebo arm (2368/4725, 50.12% versus 2357/4725, 49.88%)

The overall response rates (both participant-completed and study site-completed surveys) to the first and final Vaxtracker surveys were 99.10% (4682/4725) and 99.27% (4648/4648), respectively. The participant completion rate for both surveys was 96.47% (4558/4725) and 96.65% (4525/4682). There were no statistically significant differences in participant response rate by treatment arm for the first survey (vaccine: 2290/2368, 97.71%; placebo: 2268/2357, 96.22%; P=.37) or final survey (vaccine: 2268/2346, 96.68%; placebo: 2257/2336, 96.62%; P=.91). When reviewing response rate by gender, females were more likely to respond to the first survey compared with males (first survey: 2228/2293, 97.17% versus 2330/2432, 95.81%; P=.01). Females were also more likely to complete the final survey when compared with males; however, this difference was not statistically significant (Table 1).

The times from link dispatch (via email or SMS text message) to survey submission were calculated for participant-completed questionnaires. The median time from dispatch to main survey submission was 4 hours (range 0.05 hours-47 days), with 90% (4103/4558) of surveys completed within 71 hours of the link being sent. For the final survey, the median time between dispatch and submission was 5 hours (range 0.03 hours-74 days) with 90% (4072/4525) of surveys completed within 72 hours of dispatch.

Although the majority of participants provided both an email address and mobile number, when comparing the method by which the link was sent (via email or via SMS text message), AUSPICE Vaxtracker participants were more likely to respond via a link sent by email for both surveys and when stratifying by gender (P<.001). Males were more likely to respond via link sent by email for the final survey compared with females (1525/2323, 66.65% versus 1368/2207, 61.98%; P=.01). Males were also more likely to respond from a desktop or laptop computer for both surveys compared with females, whereas females preferred to respond via their mobile device for both surveys (Table 1).

Table 1. Response rates to the AUSPICE Vaxtracker first (N=4725) and final survey (N=4682).

Survey attributes by survey type	Male, n (%)	Female, n (%)	*P* value
First survey			
Number of people enrolled	2432 (51.47)	2293 (48.53)	—[a]
Overall response rate	2403 (98.81)	2279 (99.39)	.04
Response rate by participants	2330 (95.81)	2228 (97.17)	.01
Response rate via link type			
Email link	1440 (61.38)	1333 (60.15)	.40
SMS[b] text message link	890 (38.54)	895 (40.77)	.13
By device type			
Mobile device (mobile phone/tablet)	1151 (47.29)	1337 (58.31)	<.001
Computer (desktop/laptop)	1179 (48.48)	891 (38.86)	<.001
Final survey[c]			
Number of people eligible for final survey	2403 (51.32)	2279 (48.68)	—
Overall response rate	2387 (99.33)	2261 (99.21)	.63
Response rate by participants	2313 (96.25)	2212 (97.06)	.12
Response via link type			
Email link	1525 (66.65)	1368 (61.98)	.01
SMS text message link	788 (34.44)	844 (38.66)	.003
By device type			
Mobile device (mobile phone/tablet)	1104 (45.94)	1369 (60.07)	<.001
Computer (desktop/laptop)	1209 (50.31)	843 (33.99)	<.001

[a]Not applicable.

[b]SMS: short message service.

[c]Only participants completing the first survey, online or via a study site staff member, received the second survey.

Adverse events following immunization were reported by 34.40% (805/2340) of 23vPPV recipients in the 7 days following vaccination (Table 2). Symptoms most often reported by vaccine recipients were pain at injection site (587/2340, 25.09%), limited arm movement (224/2340, 9.57%), and generalized muscle pain (217/2340, 9.27%). In total, 240 of 2332 (10.29%) placebo recipients reported symptoms following vaccination; fatigue and headaches were the dominant symptoms reported (103/2332, 4.42%; 98/2234, 4.20%, respectively). When reviewing participants reporting symptoms by vaccine, a higher proportion of placebo recipients sought medical advice in the 7 days following vaccination compared with 23vPPV recipients (18/240, 7.5% versus 15/805, 1.9%, *P*=.001). One patient reported visiting an emergency department in the 7 days following vaccination with the placebo for an unrelated reason.

Table 2. Symptom profiles for participants receiving 23vPPV and placebo in the seven days following vaccination

Symptoms	23vPPV (N=2340), n (%)	Saline placebo (N=2332), n (%)	*P* value
Any symptom	805 (34.40)	240 (10.30)	<.001
Any local symptoms	660 (28.21)	59 (2.53)	<.001
Redness at injection site	146 (6.24)	12 (0.51)	<.001
Swelling at injection site	207 (8.85)	12 (0.51)	<.001
Pain at injection site	587 (25.09)	49 (2.10)	<.001
Limited arm movement	224 (9.57)	11 (0.47)	<.001
Any systemic symptom	438 (18.72)	180 (7.72)	<.001
Fever (reported)	81 (3.46)	54 (2.32)	.02
Fatigue	199 (8.50)	103 (4.42)	<.001
Headache	172 (7.35)	98 (4.20)	<.001
Chills	56 (2.39)	30 (1.29)	<.001
Rash	31 (1.32)	3 (0.13)	<.001
Extensive arm swelling[a]	13 (0.55)	0 (0.0)	<.001
Generalized muscle pain	217 (9.27)	64 (2.74)	<.001
Generalized joint pain	50 (2.14)	32 (1.37)	.047
Medical advice sought (symptomatic patients)[b]	15 (1.9)	18 (7.5)	<.001

[a]Described as elbow joint to shoulder joint.

[b]23vPPV: N=790, saline placebo: N=240.

When comparing AEFI rates, females were more likely to report any type of reaction for both the vaccine and placebo when compared to males (vaccine: 433/1138, 38.05% versus 372/1202, 30.95%; *P*<.001; placebo: 141/1139, 12.38% versus 99/1193, 8.30%; *P*=.001) in the 7 days following vaccination. When stratifying reaction rates into broad reaction types for vaccine recipients (any local symptoms, any systemic symptoms), the difference in symptom rates between gender still applied (Table 3). Similarly, females receiving the placebo were more likely to report any symptoms (141/1139, 12.38% versus 99/1193, 8.30%; *P*=.001) and any systemic symptoms (105/1139, 9.22% versus 75/1193, 6.29%; *P*=.008) compared with males; however, this difference was not statistically significant for the reporting of any local symptoms. When reviewing individual symptom rates, females were more likely to report most individual symptoms following vaccination compared with males (Table 3) for both the placebo and the vaccine. Females were also more likely to report extensive arm swelling from elbow to shoulder joint (10/1138, 0.88% versus 3/.25, 0.25%; *P*=.04) after vaccination with 23vPPV compared with males (Table 3), although none of the patients reporting this symptom sought medical advice in the 7 days following vaccination. Males were more likely to report the symptom of fever (not measured) for both the vaccine and the placebo; however, this difference was not significant.

Vaccine recipients were more likely to report an AEFI during an interview by a study site member compared with those who completed a survey online (26/123, 47.3% versus 779/4559, 34.09%; *P*=.04). The overall AEFI rates for placebo recipients were slightly higher for those who were interviewed by study site staff; however, this difference was not statistically significant. When stratifying AEFI rates by gender and method of completion (interviewed by study site versus completed survey online), male vaccine recipients who were interviewed by study site staff were more likely to report an AEFI compared with females and compared with both male and female placebo recipients (Table 4).

Table 3. Symptom profiles for participants receiving 23vPPV and placebo in the seven days following vaccination by gender.

Symptoms	23vPPV (n=2340)			Saline placebo (n=2332)		
	Female (n=1138), n (%)	Male (n=1202), n (%)	*P* value	Female (n=1139), n (%)	Male (n=1193), n (%)	*P* value
Any symptom	433 (38.05)	372 (30.95)	<.001	141(12.38)	99 (8.30)	.001
Any local symptoms	363 (31.90)	297 (24.71)	<.001	36 (3.16)	23 (1.93)	.06
Redness at injection site	89 (7.82)	57 (4.74)	.002	10 (0.88)	2 (0.17)	.02
Swelling at injection site	120 (10.54)	87 (7.24)	.005	10 (0.88)	2 (0.17)	.02
Pain at injection site	321 (28.21)	266 (22.13)	.001	28 (2.46)	21 (1.76)	.24
Limited arm movement	157 (13.80)	67 (5.57)	<.001	8 (0.70)	3 (0.25)	.11
Any systemic symptom	236 (20.74)	201 (16.72)	.01	105 (9.22)	75 (6.29)	.008
Fever (reported)	31 (2.72)	50 (4.16)	.06	23 (2.02)	31 (2.60)	.35
Fatigue	109 (9.58)	90 (7.49)	.07	64 (5.62)	39 (3.27)	.01
Headache	97 (8.52)	75 (6.24)	.03	63 (5.53)	35 (2.93)	.002
Chills	30 (2.64)	26 (2.16)	.45	18 (1.58)	12 (1.01)	.22
Rash	23 (2.02)	8 (0.67)	.004	1 (0.09)	2 (0.17)	.59
Extensive arm swelling[a]	10 (0.88)	3 (0.25)	.04	0 (0.00)	0 (0.00)	
Generalized muscle pain	114 (10.02)	103 (8.57)	.23	48 (4.21)	16 (1.34)	<.001
Generalized joint pain	29 (2.55)	21 (1.75)	.18	21 (1.84)	11 (0.92)	.056

[a]Described as elbow joint to shoulder joint.

Table 4. Reported symptoms following immunization by method of survey completion (participant or study site staff member) and gender.

Method, vaccine type, and gender	Participant-completed surveys (n=4549)[a], n (%)	Study site-completed surveys (n=123)[b], n (%)	*P* value
Overall AEFI[c] rate			
23vPPV recipients	779 (34.09)	26 (47.27)	.04
Saline recipients	232 (10.25)	8 (11.76)	.69
AEFI rate by gender			
Female			
23vPPV recipients	424 (37.89)	9 (47.37)	.40
Saline recipients	136 (12.3)	5 (15.6)	.57
Male			
23vPPV recipients	355 (30.45)	17 (47.22)	.03
Saline recipients	96 (8.3)	3 (8.3)	.99

[a]Denominators: F=2323, M=2226. Vaccine recipients by gender. 23V PPV: M=1166, F=1119. Saline: F=1107, M=1157.

[b]Denominators: F=51, M=72. Vaccine doses by gender. 23V PPV: M=36, F=19. Saline: F=32, M=36.

[c]AEFI: adverse events following immunization.

Discussion

Vaxtracker proved useful for monitoring adverse events in an older cohort participating in a randomized controlled trial of vaccines, with high participation rates for all study sites, age groups, and genders using both mobile devices and/or stand-alone computers. Over 90% of AUSPICE Vaxtracker participants completed both surveys within three days of receipt of the survey link, with a median time of 4 hours (0.05 hours-54 days) and 5 hours (0.03 hours-74 days) from receipt of survey link for the main and final surveys, respectively, demonstrating Vaxtracker's potential for rapidly detecting adverse events in close to "real time."

Overall, 23vPPV was well-tolerated in study participants enrolled in AUSPICE, with individual reaction rates lower than those reported by Jackson et al [10], noting the differences in reporting timeframes (7 days compared with 14 days postvaccination) and the age of the cohort. Interestingly, our

study identified a statistically significant difference in overall AEFI rates in patients who were interviewed by study site staff members when compared with participants completing their own surveys online. Stratifying by gender, this statistically significant difference was identified only in men who were interviewed by study site staff. Median survey completion times (from vaccination date to survey completion) was 7 days for responding participants and 19 days for participants interviewed by study site (noting that nonresponders were proactively followed up by study sites after two reminder messages, or 16 days after vaccination). Although this difference warrants consideration when comparing methods of survey completion and AEFI rates, the small number of participants interviewed (N=123) and apparent recall bias as a result of delays in survey submission should not be discounted. Despite these differences, Vaxtracker provides reassuring safety data should the AUSPICE study confirm the protective benefits of earlier pneumococcal vaccination to reduce cardiovascular events.

A finding from our study was that females were consistently more likely than males to report an AEFI following vaccination with 23vPPV based on individual symptoms (redness, swelling, or pain at the injection site; extensive arm swelling from joint to joint; limited arm movement; headache and rash) and broad symptom type (any reaction, any local reaction, and any systemic reaction). Higher AEFI rates in females have been reported previously, with heightened reactogenicity to vaccination likely due to biological (heightened immune response), social, and behavioral factors [13-15]. Sex differentials for patient outcome reporting and active vaccine safety surveillance require further investigation.

When reviewing the preferred method of response to the Vaxtracker surveys, our study identified that females in the age group of 55 to 61 years were slightly more likely to respond to an online questionnaire via a mobile device (mobile phone, tablet) compared with males. The preference of responding to a questionnaire link by computer may relate to the nature of employment or employment status; however, this cannot be confirmed because occupation and employment status were not collected in this study. The gender preference for a device type (computer versus mobile device) is an important finding from this study because it indicates that, for optimal response rates, online surveys should be sent to both email and SMS text message services, with questionnaires designed for both device types.

As expected, there were statistically significant differences in individual symptom reaction rates between the vaccine and placebo groups. By monitoring “adverse events following vaccination” in the blinded saline placebo arm, background symptom levels could be assessed and a true difference with the vaccine arm occurring in the community for this age group during the surveillance period could be determined. Background levels of symptoms occurring in the community need to be considered when interpreting postlicensure AEFI surveillance data.

A limitation of this study is that a formal cost evaluation of AUSPICE Vaxtracker has not yet been conducted. Descriptive analysis of the number of messages automatically dispatched to Vaxtracker participants was conducted, with approximately 17,200 emails and 16,900 SMS text messages to the 4725 people enrolled in Vaxtracker. In addition, 1350 emails noting serious symptom alerts were sent via email to AUSPICE and Vaxtracker staff. Study staff interviewed 247 participants when questionnaires had not been completed (due to no access to an email address or mobile phone, or those who had not responded to the questionnaire), and 82 participants when a serious symptom alert had been noted. These data suggest that Vaxtracker was an efficient method for monitoring AEFI in people participating in a large clinical trial. When considering the convenience of completing a survey by participants at a time suited to them, the usefulness of Vaxtracker in collecting timely AEFI data is obvious.

Postlicensure AEFI surveillance in near-real time is critical for monitoring the introduction of new vaccines into the community and providing assurance to people participating in research in which the additional benefits of a vaccine are being explored. For a participant-centered surveillance system to be successful in the detection of vaccine safety signals, the program needs to consider the target audience. Flexibility that caters for both mobile devices and stand-alone computer systems increases the system’s usability across age groups, as shown in this study. High response rates by older people participating in AUSPICE indicates a high level of acceptance of Vaxtracker and demonstrates the program’s effectiveness in monitoring vaccine safety in people aged 55 to 60 years. Finally, the use of automated email and SMS text message to send links to surveys provides a timely and convenient method of collecting data from people participating in large clinical trials.

Acknowledgments

The authors would like to thank the following people: Stephen Clarke; the AUSPICE study staff from the Adelaide, Canberra, Gosford, Melbourne, Newcastle, and Perth sites; the AUSPICE Study Team; and AUSPICE participants. The AUSPICE study is funded by a National Health and Medical Research Grant (ID#1062536).

Conflicts of Interest

None declared.

References

1. Cashman P, Macartney K, Khandaker G, King C, Gold M, Durrheim DN. Participant-centred active surveillance of adverse events following immunisation: a narrative review. Int Health 2017 May 01;9(3):164-176 [FREE Full text] [doi: 10.1093/inthealth/ihx019] [Medline: 28582563]
2. Cashman P, Moberley S, Dalton C, Stephenson J, Elvidge E, Butler M, et al. Vaxtracker: active on-line surveillance for adverse events following inactivated influenza vaccine in children. Vaccine 2014 Sep 22;32(42):5503-5508. [doi: 10.1016/j.vaccine.2014.07.061] [Medline: 25077424]
3. Cashman P, Moberley S, Durrheim D. MMRV vaccine safety. Vaccine 2019 Jul 09;37(30):3946 [FREE Full text] [doi: 10.1016/j.vaccine.2018.02.108] [Medline: 29526373]
4. Ren S, Hure A, Peel R, D'Este C, Abhayaratna W, Tonkin A, AUSPICE study group. Rationale and design of a randomized controlled trial of pneumococcal polysaccharide vaccine for prevention of cardiovascular events: The Australian Study for the Prevention through Immunization of Cardiovascular Events (AUSPICE). Am Heart J 2016 Jul;177:58-65. [doi: 10.1016/j.ahj.2016.04.003] [Medline: 27297850]
5. Pillsbury A, Cashman P, Leeb A, Regan A, Westphal D, Snelling T, AusVaxSafety, surveillance team. Real-time safety surveillance of seasonal influenza vaccines in children, Australia, 2015. Euro Surveill 2015 Oct 29;20(43):6-10 [FREE Full text] [doi: 10.2807/1560-7917.ES.2015.20.43.30050] [Medline: 26536867]
6. Musher DM, Manof SB, Liss C, McFetridge RD, Marchese RD, Bushnell B, et al. Safety and antibody response, including antibody persistence for 5 years, after primary vaccination or revaccination with pneumococcal polysaccharide vaccine in middle-aged and older adults. J Infect Dis 2010 Feb 15;201(4):516-524. [doi: 10.1086/649839] [Medline: 20092407]
7. Trinh L, Macartney K, McIntyre P, Chiu C, Dey A, Menzies R. Investigating adverse events following immunisation with pneumococcal polysaccharide vaccine using electronic General Practice data. Vaccine 2017 Mar 13;35(11):1524-1529. [doi: 10.1016/j.vaccine.2017.01.063] [Medline: 28202210]
8. Miller ER, Moro PL, Cano M, Lewis P, Bryant-Genevier M, Shimabukuro TT. Post-licensure safety surveillance of 23-valent pneumococcal polysaccharide vaccine in the Vaccine Adverse Event Reporting System (VAERS), 1990-2013. Vaccine 2016 May 27;34(25):2841-2846 [FREE Full text] [doi: 10.1016/j.vaccine.2016.04.021] [Medline: 27087150]
9. Australian Technical Advisory Group on Immunisation (ATAGI). Pneumococcal disease. In: The Australian Immunisation Handbook. Canberra, Australia: Australian Government: National Health and Medical Research Council; 2018.
10. Jackson LA, Gurtman A, van Cleeff M, Jansen KU, Jayawardene D, Devlin C, et al. Immunogenicity and safety of a 13-valent pneumococcal conjugate vaccine compared to a 23-valent pneumococcal polysaccharide vaccine in pneumococcal vaccine-naive adults. Vaccine 2013 Aug 02;31(35):3577-3584. [doi: 10.1016/j.vaccine.2013.04.085] [Medline: 23688526]
11. Hammitt LL, Bulkow LR, Singleton RJ, Nuorti JP, Hummel KB, Miernyk KM, et al. Repeat revaccination with 23-valent pneumococcal polysaccharide vaccine among adults aged 55-74 years living in Alaska: no evidence of hyporesponsiveness. Vaccine 2011 Mar 09;29(12):2287-2295. [doi: 10.1016/j.vaccine.2011.01.029] [Medline: 21255685]
12. Microsoft Corporation. ASP.NET MVC Pattern. Microsoft Data Platform URL: https://dotnet.microsoft.com/apps/aspnet/mvc and https://www.microsoft.com/en-au/sql-server/ [accessed 2018-09-06]
13. Fink AL, Klein SL. Sex and gender impact immune responses to vaccines among the elderly. Physiology (Bethesda) 2015 Nov;30(6):408-416 [FREE Full text] [doi: 10.1152/physiol.00035.2015] [Medline: 26525340]
14. Klein SL, Marriott I, Fish EN. Sex-based differences in immune function and responses to vaccination. Trans R Soc Trop Med Hyg 2015 Jan;109(1):9-15 [FREE Full text] [doi: 10.1093/trstmh/tru167] [Medline: 25573105]
15. Harris T, Nair J, Fediurek J, Deeks SL. Assessment of sex-specific differences in adverse events following immunization reporting in Ontario, 2012-15. Vaccine 2017 May 02;35(19):2600-2604 [FREE Full text] [doi: 10.1016/j.vaccine.2017.03.035] [Medline: 28365252]

Abbreviations

AEFI: adverse events following immunization
AUSPICE: Australian Study for the Prevention through Immunisation of Cardiovascular Events
SMS: short message service

Edited by G Eysenbach; submitted 23.05.19; peer-reviewed by H Demir, MY Song; comments to author 22.06.19; revised version received 28.06.19; accepted 29.06.19; published 23.10.19.

Please cite as:
Munnoch SA, Cashman P, Peel R, Attia J, Hure A, Durrheim DN
Participant-Centered Online Active Surveillance for Adverse Events Following Vaccination in a Large Clinical Trial: Feasibility and Usability Study
J Med Internet Res 2019;21(10):e14791
URL: https://www.jmir.org/2019/10/e14791
doi:10.2196/14791
PMID:31647470

Original Paper

An Online Survey for Pharmacoepidemiological Investigation (Survey of Non-Medical Use of Prescription Drugs Program): Validation Study

Joshua Curtis Black[1], PhD; Karilynn Rockhill[1], MPH; Alyssa Forber[1], MSc; Elise Amioka[1], MSc; K Patrick May[1], MSc; Colleen M Haynes[1], MPH; Nabarun Dasgupta[1,2], MPH, PhD; Richard C Dart[1], MD, PhD

[1]Rocky Mountain Poison & Drug Center, Denver Health and Hospital Authority, Denver, CO, United States

[2]Injury Prevention Research Center, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

Corresponding Author:
Joshua Curtis Black, PhD
Rocky Mountain Poison & Drug Center
Denver Health and Hospital Authority
1391 N Speer Blvd, #600, M/C 0180
Denver, CO,
United States
Phone: 1 303 389 1652
Fax: 1 303 389 1482
Email: joshua.black@rmpds.org

Abstract

Background: In rapidly changing fields such as the study of drug use, the need for accurate and timely data is paramount to properly inform policy and intervention decisions. Trends in drug use can change rapidly by month, and using study designs with flexible modules could present advantages. Timely data from online panels can inform proactive interventions against emerging trends, leading to a faster public response. However, threats to validity from using online panels must be addressed to create accurate estimates.

Objective: The objective of this study was to demonstrate a comprehensive methodological approach that optimizes a nonprobability, online opt-in sample to provide timely, accurate national estimates on prevalence of drug use.

Methods: The Survey of Non-Medical Use of Prescription Drugs Program from the Researched Abuse, Diversion and Addiction Related Surveillance (RADARS) System is an online, cross-sectional survey on drug use in the United States, and several best practices were implemented. To optimize final estimates, two best practices were investigated in detail: exclusion of respondents showing careless or improbable responding patterns and calibration of weights. The approach in this work was to cumulatively implement each method, which improved key estimates during the third quarter 2018 survey launch. Cutoffs for five exclusion criteria were tested. Using a series of benchmarks, average relative bias and changes in bias were calculated for 33 different weighting variable combinations.

Results: There were 148,274 invitations sent to panelists, with 40,021 who initiated the survey (26.99%). After eligibility assessment, 20.23% (29,998/148,274) of the completed questionnaires were available for analysis. A total of 0.52% (157/29,998) of respondents were excluded based on careless or improbable responses; however, these exclusions had larger impacts on lower volume drugs. Number of exclusions applied were negatively correlated to total dispensing volume by drug (Spearman ρ=–.88, P<.001). A weighting scheme including three demographic and two health characteristics reduced average relative bias by 31.2%. After weighting, estimates of drug use decreased, reflecting a weighted sample that had healthier benchmarks than the unweighted sample.

Conclusions: Our study illustrates a new approach to using nonprobability online panels to achieve national prevalence estimates for drug abuse. We were able to overcome challenges with using nonprobability internet samples, including misclassification due to improbable responses. Final drug use and health estimates demonstrated concurrent validity to national probability-based drug use and health surveys. Inclusion of multiple best practices cumulatively improved the estimates generated. This method can bridge the information gap when there is a need for prompt, accurate national data.

(J Med Internet Res 2019;21(10):e15830) doi:10.2196/15830

KEYWORDS

nonprobability methods; general population survey; drug abuse; calibration weights

Introduction

Large governmental surveys, such as the National Survey on Drug Use and Health (NSDUH) in the United States, are used for nationwide drug use surveillance, offering researchers substantial statistical power for subgroup analyses, questionnaire consistency over decades, comprehensive and validated questionnaires, and probability-based geographic sampling for nationally representative estimates. However, these types of surveys cost millions of dollars a year to conduct, require training of field agents, and have a 2-year lag for data publication [1].

Trends in drug use can change rapidly by month, and using study designs with flexible modules could present advantages. Timely data can inform proactive interventions against emerging trends, leading to a faster public response. Large population surveys have used computer-assisted interviewing [2-4] with increased accuracy of self-reported socially stigmatized behaviors [5,6]. Internet-based questionnaires are an extension of computer-assisted interviewing, albeit with additional sampling and validity concerns, but van Gelder et al [7] have specifically suggested that illicit drug use may be a use case for internet-based questionnaires in epidemiology [7].

The use of online panels for public health research has grown in recent years [8-11]. Survey panels are groups of individuals who opt in to take surveys for modest compensation on a wide array of topics, maintained by commercial panel-access vendors. The sampling frame is theoretically suitable, since 90% of US adults use the internet [12]. Beyond efficient and rapid recruitment, panels offer superior anonymity and reductions in social desirability bias compared with in-person interviews [8,9].

However, threats to validity unique to internet surveys require removing careless or improbable responses [13], calibrating sample representativeness [14], preventing missing data [15], and addressing low response rates [16]. Crucially, representativeness of the sample to the target population requires methodological development since straightforward approaches, like poststratification demographic weighting, are insufficient [17].

This paper describes the development of a comprehensive methodology that addresses threats to validity of using survey panels for national drug use estimates. The approach encompasses mobile-friendly interface, skip logic [18], response randomization [19], careless/improbable response exclusions [13,20], and calibration weighting [21,22]. External validity was assessed compared with three probability-based national surveys. To our knowledge, this is the first use of online panel data incorporating multiple best practices to produce nationally valid estimates regarding drug use.

Methods

Survey Overview

The Researched Abuse, Diversion, and Addiction-Related Surveillance (RADARS) System comprises multiple data sources that characterize and monitor drug use [23]. The goal of the Survey of Non-Medical Use of Prescription Drugs (NMURx) Program described here is to provide accurate and timely estimates of prescription drug nonmedical use (NMU) and associated motivations and behaviors in the adult general US population. The NMURx Program employs a cross-sectional, opt-in online self-administered questionnaire drawn from a commercial survey panel. Respondents' personal information is kept confidential by the survey administrators; personally identifiable information is not collected on the questionnaire, and information held by the survey administrators is not available to the researchers. Following best practices for implementation of online questionnaires [19], three methodological practices are described: reduction of order effect bias by randomization of question order, exclusion criteria based on careless/improbable responses, and calibration weighting for generalizability.

Questionnaire Development

The main body of the questionnaire covered motivations and behaviors surrounding prescription drug use of four prescription drug classes (pain relievers, sedatives, stimulants, and cannabinoids), documenting lifetime and last 12-month NMU. NMU of prescription drugs was defined as use "in a way not directed by your health care provider." Examples of the questionnaire questions can be found in Multimedia Appendix 1, Section A, followed by a list of drug classes and substances included in the questionnaire (Multimedia Appendix 1, Section B). Additional sections in the survey (some not included in this analysis) are: demographics, Drug Abuse Screening Test (DAST-10) measuring severity of problematic drug use [24], motivations and drug use behavior (eg, reasons for use, route of administration, source of acquisition), and health status (eg, substance use disorder treatment history, mental health disorders). Skip logic was used to minimize the number of questions a respondent was required to answer, with focus on preventing motivated underreporting [18]. Due to the large number of drugs included on the questionnaire and the follow-up questions on behaviors, motivated underreporting was of particular concern. The Checklist for Reporting Results of Internet E-Surveys (CHERRIES) is provided for further details on survey development and implementation (Multimedia Appendix 1, Section C).

Two strategies were used to reduce order effect bias. First, the order of drug classes was randomized, followed by order of substances within each drug class. Block randomization kept together similar substances (eg, all pain relievers), with the order consistently maintained throughout the survey sections. Second, respondents were forced to provide product-specific answers for last 12-month NMU of specific drug products that

had been endorsed at the class level. This was intended to improve internal validity and further reduce order effect bias [19].

Participant Sampling

The survey was open from September 28 through November 21, 2018. To be eligible for the survey, respondents must have been aged 18 years or older, and they must not have completed a NMURx Program survey in the same calendar year, excluding a small number of potential respondents who participated in pilot surveys. The panel company recruited panelists from the US population without specific consideration for the NMURx Program survey; selection into the panel was nonprobability-based and was a self-selected population of people who take surveys for compensation. The panel company employs evolving proprietary techniques to ensure panelists are providing reliable responses, with inactive or fraudulent accounts culled regularly. The panel company calculated each panelist's expected response likelihood based on recent activity. These probabilities were used to select a random sample of panelists expected to yield 30,000 completed questionnaires. The email invitation did not include information about the survey topic to minimize selection bias; the topic was provided once a panelist opened the link during the consenting process.

To reduce the possibility of extreme analytical weights, 8-stratum sampling quotas were devised, proportionally based on the adult residential population from the American Community Survey (ACS) [25], stratified by male/female gender for four Census regions (Northeast, Midwest, South, and West). Based on pilot experience, each stratum was allowed a –25% and +10% range of acceptable number of surveys. Once all quotas reached their minimum and at least 30,000 questionnaires were completed, the survey link was closed. Survey respondents were compensated roughly US $1.

Exclusion Criteria Assessment

Due to programmable internal data consistency checks, an outstanding concern after survey administration was the identification of completed questionnaires exhibiting careless or improbable response patterns (eg, endorsing all drugs at biologically improbable frequency). Methods used to exclude responses were adapted [13,20] to generate exclusion criteria that were validated against other questionnaire elements.

Based on previous implementation experience and literature review of consumer product surveys, seven different metrics were investigated as possible exclusion criterion using multiple hurdles [13], and four were chosen based on performance: (1) consecutive positive use endorsements of up to 42 prescription drugs based on the LongString method [20], (2) alternating patterns of yes/no endorsement of prescription drugs based on the even-odd consistency method [20] with Pearson correlations no more negative than –0.6, (3) alternating patterns of illicitly manufactured drugs with Pearson correlations no more negative than –0.8 for fewer drugs, and (4) total number of specific products endorsed for NMU in last 12 months via modified outlier analysis (out of 298 possible, most respondents only endorsed a handful) [13]. Completion time of 8 introductory questions in less than 16 seconds [13] did not provide additional discriminatory value (data not shown).

Since no gold standard was available for validation of the critical lifetime and 12-month prevalence questions, three internal consistency metrics were developed. Derived from other survey sections, these metrics provided biologically plausible support for responses to NMU: (1) survey response time for lifetime prescription drug use was a proxy for completion speed; (2) Mahalanobis distances were calculated on lifetime prescription, nonprescription, and illicitly manufactured drug use responses, representing deviance compared to the entire sample [13]; and (3) proportion of contradictory answers was calculated. For example, respondents were asked the time frame in which they had first initiated NMU and when they most recently nonmedically used, and the skip logic of the questionnaire allows for contradictory answers. Cut points were identified based on established theories (described in Results), visual inspection, and correlation coefficients [26].

To evaluate internal consistency, cut points were evaluated against demographics, drug use behaviors, and overall drug endorsements. To evaluate external consistency for careless responses, we compared relative endorsements to national opioid dispensing data from the US-based Longitudinal Patient Databases (IQVIA Inc), a standard source that provides estimates of dosage units dispensed in retail pharmacies. Since low-volume drugs should result in fewer endorsements, we hypothesized that careless responses would be roughly proportional to dispensing (using Spearman correlation), and excluded responses would account for a larger proportion of low-volume drugs.

Calibration Weighting for National Estimates

A calibration weighting scheme was developed to generate national estimates for the adult population. The goal of the weighting scheme was to reduce the bias in estimates resulting from the self-selection of survey panelists by forcing the distribution of our sample to look similar to national estimates across demographics and health-related variables. Generalized raking using auxiliary information with incomplete stratification was selected as the method of calibration weighting [21,22] because raking has been shown to be equivalent or superior to propensity score methods or sample matching in reducing bias [27,28]. Briefly, this method matches the marginal distributions of each variable in the sample to the marginal distribution from the population by iteratively adjusting the base weights. The base weight (w_b) was calculated where N is the adult population in 2017 (N=252,063,800) and n is the sample size (Figure 1).

Figure 1. Base weight equation.

$$w_b = \frac{N}{n} \quad (1)$$

The analytical weight was calculated using established procedures [29]. Maximum tolerance was 0.1 percentage points; convergence occurred at a tolerance of 0.1 weighted frequency. The national marginal values were obtained from two 2017 probability-based national surveys, ACS and National Health Interview Survey (NHIS) [2,25].

Eight potential weighting variables from ACS and NHIS were selected based on associations with three lifetime measures also appearing in our questionnaire: any illicit drug use, any prescription pain reliever NMU, and any prescription NMU. Three demographic variables (age, sex, Census region of residence) and two household characteristics (household income and number of people in the home) were derived from ACS. Three health-related characteristics (self-assessment of general health, limitations in daily activities, and smoking tobacco) came from NHIS. To match basic demographics of the adult population, the three demographic variables were included in every model. Raking was tested against all remaining combinations (33 possible schemes).

To evaluate the 33 possible schemes, 26 benchmark national estimates were compared between the NMURx Program and four national surveys: ACS, NHIS, NSDUH, and National Health and Nutrition Examination Survey [3,4]. The absolute value of the relative difference (D_i) for the i^{th} benchmark between the NMURx Program estimate (p_i) and the national survey estimate (π_i) were calculated (Figure 2, equation 2). These were averaged across the 26 benchmark estimates (Figure 2, equation 3, where b_n is the number of benchmarks).

Figure 2. Absolute value of the relative difference and average relative difference equations.

$$D_i = \frac{|\pi_i - p_i|}{\pi_i} * 100 \quad (2)$$

$$D = \frac{\left(\sum_{i=1}^{b_n} \frac{|\pi_i - p_i|}{\pi_i}\right)}{b_n} * 100 \quad (3)$$

The average relative difference in estimates across the weighting schemes for health-related benchmark estimates (overnight stay in a hospital, pain reliever use, illicit drug use, and alcohol use) compared with nonhealth-related benchmarks (race/ethnicity, marital status, education, employment, and insurance status) was evaluated. Final analytical weights represented the number of adults that a survey respondent would represent in the United States, generating national prevalence estimates, with 95% confidence intervals using variance estimation through Taylor series linearization [30].

Ethics Review

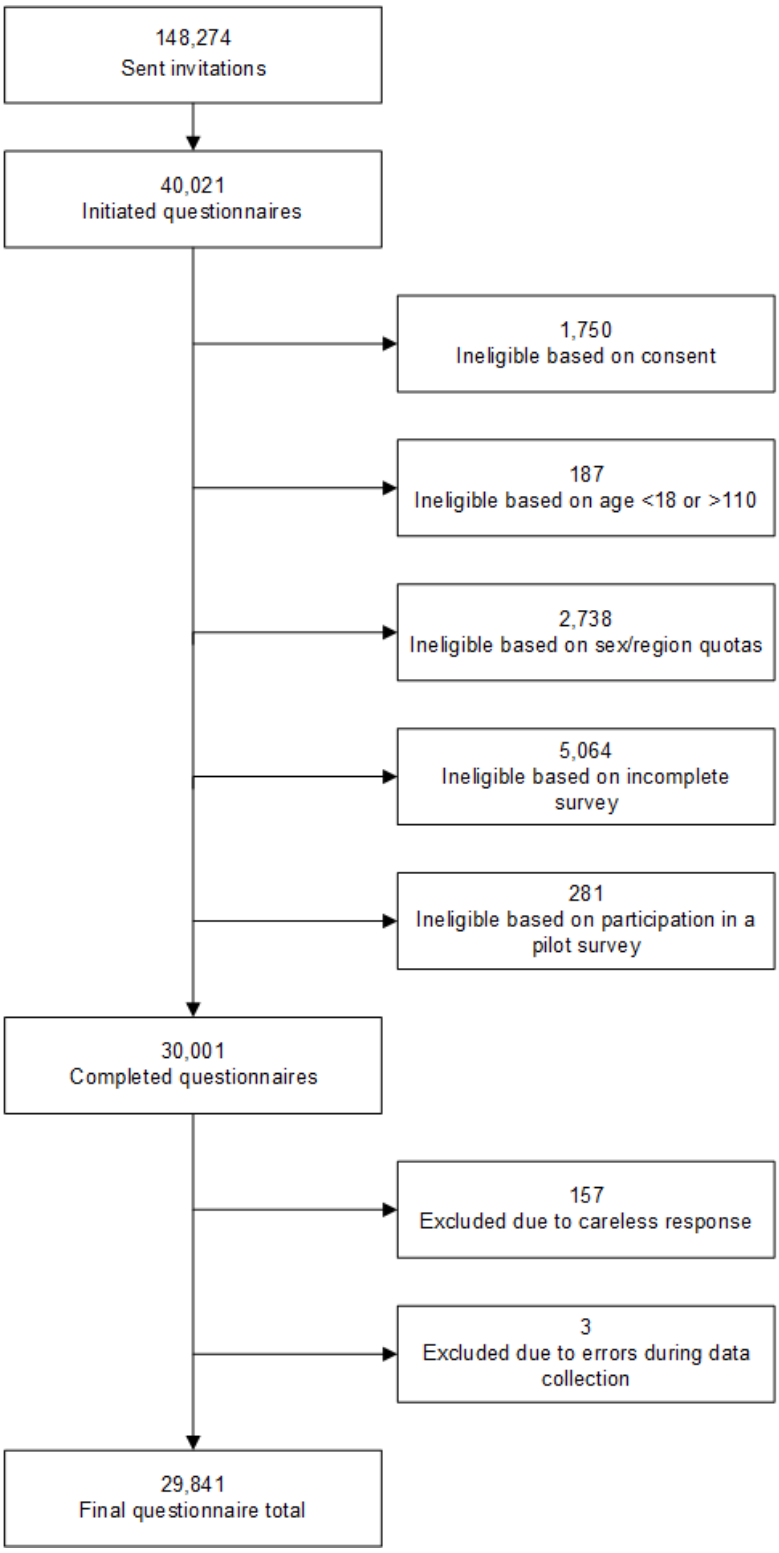
The protocol and survey instrument were reviewed and approved by the Colorado Multiple Institutional Review Board; a certificate of exemption was granted on July 5, 2016.

Results

Participant Recruitment

There were 148,274 invitations sent to panelists, and 40,021 (26.99%) people initiated the survey. After eligibility assessment, 74.96% (29,998/40,021) of the completed questionnaires were available for analysis (Figure 3). After careless responses were removed, the final participation rate was 20.13% (29,841/148,274). Out of 910 3-digit zip codes in the United States, 883 zip codes had at least one respondent. An order effect was present, and the likelihood of endorsement for individual active pharmaceutical ingredients was associated with the position in which the item was presented in their questionnaire (Multimedia Appendix 1, Section D).

Figure 3. Flow diagram of respondents invited to the survey.



Evaluations of Exclusion Criteria

There were 157 respondents (0.52% of the sample) excluded based on careless or improbable responses (Multimedia Appendix 1, Section E). The final sample had a median completion time of 10 minutes, 40 seconds. For criterion 1, only 27 responses were identified, where over half of the 42 drug use questions were consecutively endorsed (Table 1). Consistent with recommendations [13], this cut point requires a respondent to endorse at least two separate drug classes in an unbroken string of "Yes" responses to be excluded. Criteria 2 and 3 for alternating responses resulted in 33 and 17 surveys being excluded, respectively. For criterion 4 on total drug endorsements, 91.01% (27,301/29,998) did not endorse any NMU in the last 12 months. Among those endorsing at least one product, median products endorsed was 3 (interquartile range 1 to 7) of 298 possible. Given the highly skewed distribution, visual inspection of Mahalanobis distance, question

completion time, and contradictory answers were used to select 35 products in the last 12 months endorsed as a conservative cut point (Multimedia Appendix 1, Section F), resulting in 96 responses excluded. There was very little overlap across multiple criteria. There were 6 respondents who were identified by both criterion 1 and criterion 4. A total of 10 respondents were identified by both criterion 2 and criterion 3. Table 1 demonstrates that groups of respondents excluded by each criterion also demonstrated other behaviors indicative of careless response. Excluded respondents answered questions more quickly. Excluded respondents had greater mean Mahalanobis distances, and the proportion of excluded respondents who provided at least one contradictory answer on the survey was very different from the proportion of included respondents who provided contradictory answers (67/157, 42.7%, and 224/29,841, 0.75%, respectively). Excluded respondents more frequently reported being male, younger, and Hispanic compared with respondents who were not excluded, although statistical tests of differences were not conducted.

While a small proportion of responses were excluded (157/29,998, 0.52%), these exclusions had a larger impact on unweighted endorsements of lower volume active pharmaceutical ingredients (Table 2). The relative percentage decrease in responses endorsing NMU of opioid ingredients was negatively correlated to dispensing volume (Spearman ρ=–.88, P<.001), confirming our hypothesis that misclassification due to careless response would have a greater impact on low-volume products.

Table 1. Respondent characteristics of excluded and included respondents.

Characteristics	Criterion 1 (n=27)	Criterion 2 (n=33)	Criterion 3 (n=17)	Criterion 4 (n=96)	All excluded respondents[a] (n=157)	Included respondents (n=29,841)
Male, n (%)	21 (77.78)	25 (75.76)	13 (76.47)	75 (78.13)	124 (78.98)	16,065 (53.84)
Age in years, median (IQR[b])	34 (28, 37)	32 (28, 37)	35 (29, 43)	33 (28, 37)	33 (28, 37)	53 (35, 66)
Race/ethnicity[c], n (%)						
Hispanic/Latino(a)	8 (29.63)	12 (36.36)	8 (47.06)	26 (27.08)	47 (29.94)	2271 (7.61)
White	19 (70.37)	26 (78.79)	13 (76.47)	72 (75.00)	116 (73.89)	24,946 (83.60)
African American	__[d]	5 (15.15)	__[d]	17 (17.71)	27 (17.20)	2838 (9.51)
Asian	__[d]	__[d]	__[d]	__[d]	6 (3.82)	1164 (3.90)
Other	__[d]	__[d]	__[d]	__[d]	10 (6.37)	1649 (5.53)
Total time spent on prescription drug use question (seconds), median (IQR)	106.06 (80.98, 172.65)	73.26 (66.16, 89.73)	82.48 (66.16, 98.60)	97.77 (77.05, 143.40)	93.99 (73.22, 127.81)	116.31 (89.66, 159.53)
Mahalonbis distance, mean (SD)	18.89 (3.67)	28.45 (2.35)	29.28 (3.15)	24.15 (4.76)	24.53 (5.05)	6.60 (4.75)
Contradictory answers, n (%)	8 (29.63)	15 (45.45)	4 (23.53)	46 (47.92)	67 (42.68)	224 (0.75)

[a]Excluded respondents identified as any one of the four criteria established for careless response: LongString prescription drug use endorsements (criterion 1), even-odd consistency for prescription drug use (criterion 2), illicit drug use (criterion 3), or total product endorsement (criterion 4).

[b]IQR: interquartile range.

[c]Respondents may select multiple races, so percentage may not sum to 100.

[d]Cells with fewer than 5 respondents are suppressed for disclosure protections.

Table 2. Relative decrease in prescription drug nonmedical use endorsements after exclusions compared with drug availability.

Active pharmaceutical ingredient	No exclusions applied (n)	All exclusions applied (n)	Relative decrease (%)	Dispensing volume (dosage units dispensed)
Hydrocodone	905	855	5.52	4,570,914,825
Oxycodone	762	708	7.09	3,373,604,063
Tramadol	485	434	10.52	2,403,511,798
Codeine	742	687	7.41	1,986,127,916
Morphine	320	264	17.50	467,226,515
Hydromorphone	157	114	27.39	184,212,536
Tapentadol	83	53	36.14	48,107,328
Fentanyl	187	137	26.74	36,317,430
Oxymorphone	150	111	26.00	34,595,913
Dihydrocodeine	78	47	39.74	1,889,020

Selection of Weighting Scheme

The remaining 29,841 surveys were used for calibration weighting. The unweighted NMURx Program data had an average relative difference of 36.1% compared with weighted estimates (Table 3). Across the 33 weighting schemes, the addition variables resulted in decreases in the average relative difference while relative standard error increased (Multimedia Appendix 1, Section G). Little variation in average relative difference was observed among nonhealth-related benchmarks; however, there were large changes in average difference among health-related benchmarks (Figure 4). Five-variable weighting schemes appeared to maximize reduction in relative difference and with minor increases in relative standard error, resulting in the selection including age, gender, region, limitation in daily activities, and tobacco use. This scheme had a 31.2% reduction in average relative difference compared with unweighted, resulting in 381 unique weights, none of which appeared extreme. The median weight was 7782.2 (interquartile range 4690.2 to 12,662.9).

Table 3. Relative difference in benchmark estimates.

Characteristic	Survey of Non-Medical Use of Prescription Drugs Program				National survey benchmark, % (95% CI)
	Unweighted		Weighted		
	Estimate, n (%)	Relative difference[a] (%)	Estimate, % (95% CI)	Relative difference[a] (%)	
Nonhealth benchmarks					
Race/ethnicity[b,c]					
Hispanic/Latino(a)	2271 (7.61)	–52.39	8.19 (7.82-8.56)	–48.78	15.99 (15.92-16.05)
White	24,946 (83.60)	10.14	82.20 (81.69-82.71)	8.30	75.90 (75.83-75.97)
African American	2838 (9.51)	–27.66	9.94 (9.55-10.34)	–24.35	13.15 (13.09-13.21)
Asian	1164 (3.90)	–40.40	4.70 (4.41-4.99)	–28.16	6.54 (6.51-6.58)
American Indian or Alaska Native	538 (1.80)	13.81	1.61 (1.45-1.77)	1.61	1.58 (1.56-1.60)
Native Hawaiian or other Pacific Islander	108 (0.36)	–5.11	0.37 (0.29-0.45)	–3.31	0.38 (0.37-0.39)
Other	1047 (3.51)	–29.90	3.88 (3.62-4.15)	–22.43	5.01 (4.97-5.04)
Marital status[c]					
Married	15,640 (52.41)	4.45	51.93 (51.29-52.57)	3.49	50.18 (50.10-50.26)
Widowed	1928 (6.46)	7.60	5.17 (4.92-5.43)	–13.82	6.00 (5.97-6.04)
Divorced	3969 (13.30)	16.08	11.50 (11.12-11.89)	0.40	11.46 (11.41-11.51)
Separated[d]	576 (1.93)	—	1.76 (1.60-1.93)	—	2.02 (1.99-2.04)
Never married	7728 (25.90)	–14.65	29.63 (29.02-30.24)	–2.35	30.34 (30.27-30.42)
Education[c]					
Less than high school	834 (2.79)	–76.84	2.69 (2.48-2.90)	–77.73	12.07 (12.01-12.12)
High school graduate or GED[e]	5846 (19.59)	–29.12	18.65 (18.16-19.15)	–32.50	27.64 (27.57-27.71)
Some college or associate's degree	10,003 (33.52)	8.69	32.98 (32.38-33.58)	6.95	30.84 (30.77-30.91)
Bachelor's or higher degree or trade school[d]	13,158 (44.09)	N/A[f]	45.68 (45.04-46.31)	N/A	29.45 (29.38-29.52)
Employed last week[c]					
Yes	12,005 (40.23)	–34.89	45.77 (45.13-46.41)	–25.92	61.78 (61.70-61.86)
No[d]	17,836 (59.77)	—	54.23 (53.59-54.87)	—	38.22 (38.14-38.30)
Private health insurance[g]					
Yes	19,007 (63.69)	–2.25	64.81 (64.20-65.42)	–0.54	65.16 (64.21-66.11)
No[d]	10,834 (36.31)	—	35.19 (34.58-35.80)	—	34.84 (33.89-35.79)
Health benchmarks					
Overnight stay in hospital in last year[g]					
Yes	3609 (12.09)	43.13	9.65 (9.30-10.01)	14.26	8.45 (8.17-8.72)
No[d]	26,232 (87.91)	—	90.35 (89.99-90.70)	—	91.55 (91.28-91.83)
Alcohol use in past 12 months[g]					
Yes	15,554 (52.12)	–3.88	52.10 (51.46-52.74)	–3.92	54.23 (53.15-55.31)
No[d]	14,287 (47.88)	—	47.90 (47.26-48.54)	—	45.77 (44.69-46.85)

Characteristic	Survey of Non-Medical Use of Prescription Drugs Program				National survey benchmark, % (95% CI)
	Unweighted		Weighted		
	Estimate, n (%)	Relative difference[a] (%)	Estimate, % (95% CI)	Relative difference[a] (%)	
Pain reliever use[h]					
Lifetime use	18,287 (61.28)	–5.55	57.20 (56.57-57.84)	–11.83	64.90 (64.23-65.57)
Past year use	9607 (32.19)	–9.54	27.32 (26.77-27.87)	–23.24	34.88 (34.21-35.55)
Lifetime NMU[i]	3928 (13.16)	24.05	11.13 (10.74-11.51)	4.87	10.45 (10.05-10.84)
Past year NMU	2313 (7.75)	76.58	6.18 (5.89-6.47)	40.78	4.06 (3.82-4.30)
Past month NMU	1222 (4.10)	218.82	2.96 (2.77-3.15)	130.22	1.21 (1.07-1.34)
Illicit drug use[h]					
Lifetime use	6230 (20.88)	–16.70	17.53 (17.06-17.99)	–30.07	25.45 (24.85-26.05)
Last year use	1756 (5.88)	54.63	4.74 (4.49-5.00)	24.61	4.08 (3.85-4.31)
Last month use	970 (3.25)	111.67	2.45 (2.27-2.63)	59.31	1.68 (1.53-1.84)

[a]Relative difference was calculated using more significant figures than presented; due to rounding these results may appear different than calculating using estimates presented in this table.

[b]Multiple races may be selected so estimates may not sum to 100%.

[c]ACS: American Community Survey.

[d]Levels of estimates that were not included in average relative difference calculation.

[e]GED: General Educational Development test.

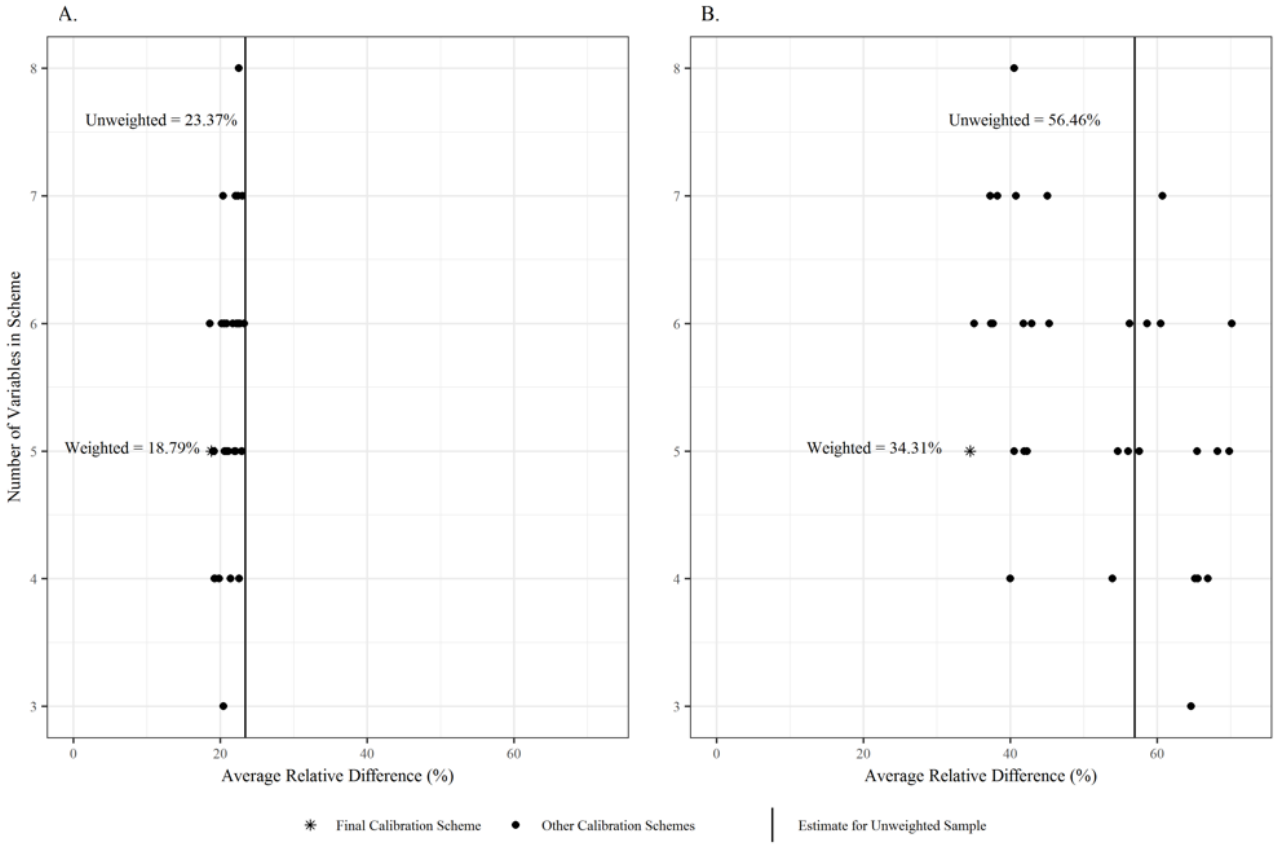
[f]Not applicable.

[g]NHIS: National Health Interview Survey.

[h]NSDUH: National Survey on Drug Use and Health.

[i]NMU: nonmedical use.

Figure 4. Average relative difference in nonhealth-related (A) and health-related (B) benchmarks with calibration weighting are shown for the 33 weighting schemes.



External Validation Results

After weighting, the NMURx Program estimates were closely aligned with the national benchmarks on demographic and other characteristics (Table 4). The unweighted sample had an overrepresentation of older adults, males, higher education, and lower household incomes, possibly a reflection of internet panel participants in general. The health profile of the sample was more similar to the national estimates after weighting. Proportions of good or excellent self-assessed health status and private insurance coverage increased while DAST-10 scores and the estimated proportion with chronic pain in the last year decreased. After weighting, NMU of any pain reliever decreased from 7.8% to 6.2%, sedatives decreased from 4.5% to 3.4%, and stimulants decreased from 2.4% to 2.0%. In addition, when comparing drug use and health indicator estimates across multiple probability surveys with similar questions, there was variation in estimates across probability-based national surveys, and NMURx Program weighted estimates were within similar ranges to the national surveys (Figure 5). Weighted estimates were closer to estimates from probability-based national surveys. Between national surveys, estimates of sex, age, and race were similar and confidence intervals generally overlapped; estimates of education varied slightly and in many cases confidence intervals didn't overlap (Multimedia Appendix 1, Section H).

Table 4. Characteristics and national estimates before and after weighting.

Characteristics	NMURx[a] Program unweighted n=29,841 n (%)	NMURx Program weighted n=252,063,800 % (95% CI)	ACS[b] weighted n=252,155,280 % (95% CI)
Age in years[c]			
18-24	2559 (8.58)	12.23 (11.75-12.71)	12.23 (12.18-12.29)
25-34	4603 (15.43)	17.80 (17.28-18.31)	17.80 (17.73-17.86)
35-44	4467 (14.97)	16.39 (15.90-16.88)	16.39 (16.33-16.45)
45-54	4161 (13.94)	16.77 (16.27-17.28)	16.77 (16.71-16.83)
55-64	5691 (19.07)	16.66 (16.22-17.10)	16.66 (16.61-16.72)
65 or more	8360 (28.02)	20.15 (19.71-20.59)	20.15 (20.09-20.21)
Sex[c]			
Male	16,065 (53.84)	48.67 (48.03-49.31)	48.67 (48.59-48.75)
Female	13,776 (46.16)	51.33 (50.69-51.97)	51.33 (51.25-51.41)
US Census region[c]			
Northeast	5219 (17.49)	17.74 (17.25-18.23)	17.74 (17.68-17.80)
Midwest	6485 (21.73)	20.90 (20.39-21.41)	20.90 (20.83-20.96)
South	11,485 (38.49)	37.74 (37.12-38.36)	37.74 (37.66-37.82)
West	6652 (22.29)	23.62 (23.06-24.17)	23.62 (23.55-23.68)
Annual household income			
<$25,000	6166 (20.66)	19.35 (18.84-19.85)	14.03 (13.98-14.09)
$25,000-$49,999	8426 (28.24)	27.17 (26.61-27.74)	19.00 (18.94-19.07)
$50,000-$74,999	6545 (21.93)	22.39 (21.86-22.93)	17.49 (17.43-17.55)
$75,000-$99,999	4070 (13.64)	14.17 (13.72-14.62)	13.60 (13.54-13.65)
$100,000 or more	4634 (15.53)	16.92 (16.43-17.41)	32.75 (32.67-32.82)
DAST-10[d] score			
None reported, 0	18,378 (61.59)	63.36 (62.74-63.98)	—[e]
Low level, 1-2	9525 (31.92)	31.48 (30.89-32.08)	—
Moderate level, 3-5	1396 (4.68)	3.83 (3.60-4.07)	—
Substantial level, 6-8	428 (1.43)	1.04 (0.92-1.15)	—
Severe level, 9-10	114 (0.38)	0.29 (0.23-0.35)	—
Self-assessed health status			
Poor	754 (2.53)	1.63 (1.49-1.76)	—
Fair	4496 (15.07)	11.48 (11.11-11.86)	—
Good	11,313 (37.91)	37.10 (36.48-37.72)	—
Very good	10,162 (34.05)	37.98 (37.35-38.61)	—
Excellent	3116 (10.44)	11.81 (11.38-12.24)	—
Chronic pain in last 12 months			
Yes	20,279 (67.96)	74.28 (73.75-74.81)	—
No	9562 (32.04)	25.72 (25.19-26.25)	—

[a]NMURx: Survey of Non-Medical Use of Prescription Drugs Program.

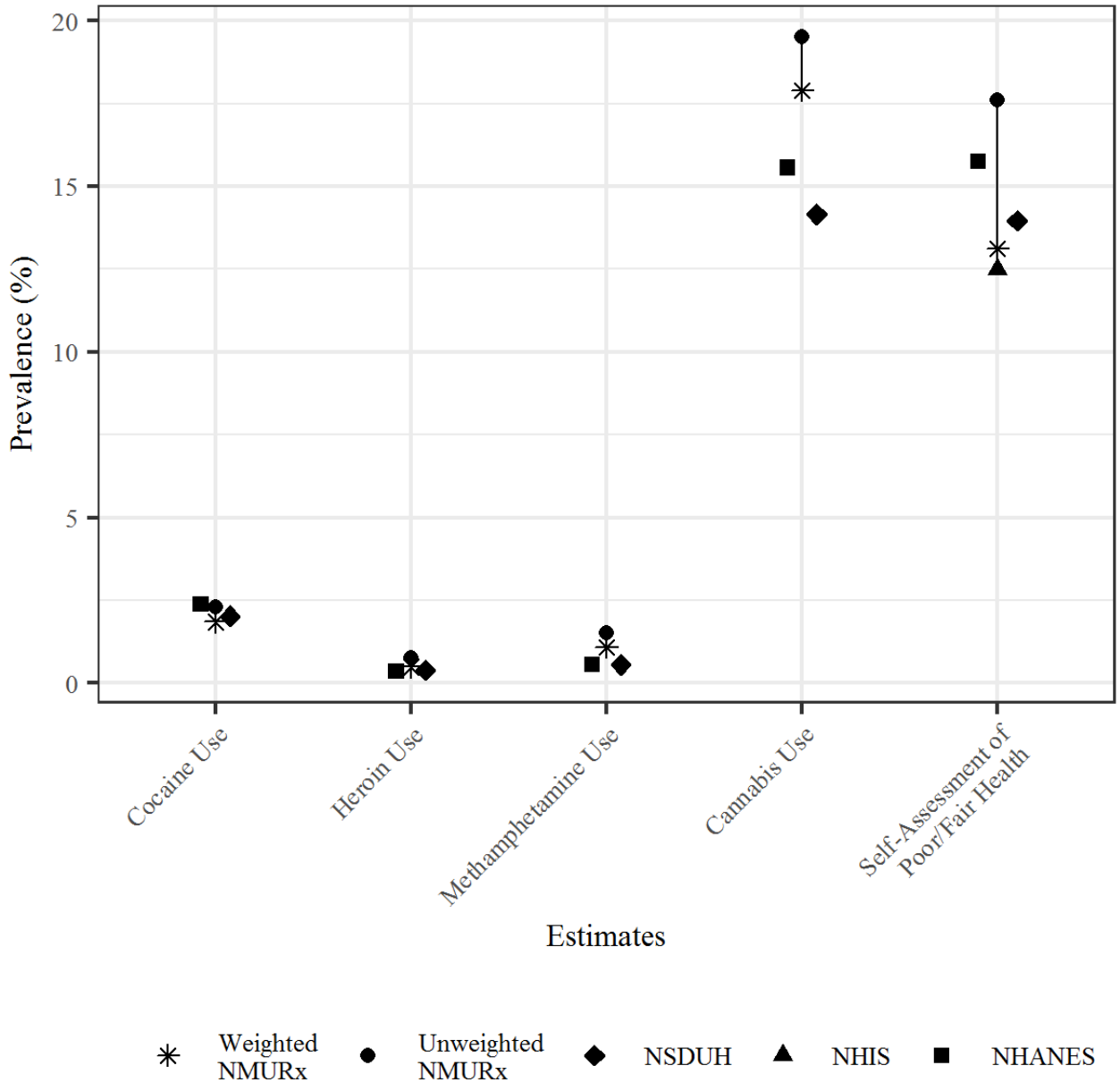
[b]ACS: American Community Survey.

[c]These variables were used in weighting scheme, so marginal estimates will align with the ACS.

[d]DAST-10: 10-item Drug Abuse Screening Test.

[e]Not applicable.

Figure 5. Comparison of estimates available across national surveys. NMURx: Survey of Non-Medical Use of Prescription; NSDUH: National Survey on Drug Use and Health; NHIS: National Health Interview Survey; NHANES: National Health and Nutrition Examination Survey.



Discussion

Principal Findings

While use of internet-based questionnaires for epidemiologic research has been previously described [31-33], our study illustrates a new approach to using nonprobability online panels to achieve national prevalence estimates for drug abuse. We were able to overcome challenges with using nonprobability internet samples [17,27,34,35], including misclassification due to careless or improbable responses. External validity of reweighted survey data demonstrated concurrent performance compared with large national probability surveys on demographics, health indicators, and drug use.

The value of internet samples is increasingly recognized [7,27,35,36], and our approach has strengths that may be relevant to drug use surveillance. Using calibration weights derived from two independent probability-sampled studies provided a hedge against overfitting [28,37]. The survey was fielded over the course of 8 weeks collecting at least 30,000 unique responses at a fraction of the cost of national probability samples. The entire process from fielding the survey to national estimates takes about 6 weeks. The ability to rapidly and inexpensively add new drugs to the survey is a considerable benefit against the background of the opioid crisis, which has evolved into its third phase, characterized by heroin-fentanyl deaths [38]. Beyond opioids, new drugs of abuse are being documented (eg, tianeptine, kratom) [39,40], while problematic drugs of the past are resurging (eg, methamphetamine, cocaine) [41]. Noncontrolled (nonscheduled) prescription drugs with abuse potential such as antidepressants [42], anticonvulsants [43], and novel psychoactive substances are not currently tracked

on national probability surveys but could easily be added to online questionnaires. Emerging novel routes of administration (eg, intra-arterial injection), fluctuations in infectious disease risk factors, and uptake of harm reduction strategies could be queried in-depth. Our results also suggest that randomization is useful in mitigating order effects on surveys and skip logic is required to prevent motivated underreporting, neither of which is common practice yet on many national surveys. The method presented here cannot replace traditional probability-based surveys; in fact, it intentionally relies on those surveys to create optimized estimates. But this method can bridge the information gap when there is a need for prompt, accurate national data.

Limitations

Ostensibly, the online-only setting creates the perception of anonymity between the respondent and researcher and reduces interviewer bias, but the role of the panel company as an intermediary and fears of data breaches may exert selection bias. There are putative gaps in the sampling frame since not all US adults use the internet. In terms of precision, 95% confidence intervals do not represent true 95% coverage probabilities because the exact selection probability from the sampling frame into the sample is not known, limiting statistical inferences within a purely frequentist context. Rather, the confidence intervals demonstrate precision of the estimates within the sampling framework, and inferences are useful when combined with an understanding of how the sample was obtained and weighted. Finally, a nonresponse adjustment was not included in this method. A drawback of using online panels is that information on nonresponding panelists is not available, and future extension of this work will be to obtain sufficient information in other ways to address this.

Conclusions

We describe a practical approach to providing a timely perspective on drug abuse in the United States, with results obtained within 6 weeks of questionnaire deployment. The approach presented mitigates many valid concerns about the use of nonprobability internet panels and could be of use to other subject domains.

Acknowledgments

The authors thank Dr Lori A Crane from the University of Colorado, Colorado School of Public Health, for her consultation on question development and questionnaire structure and Joshua Swanson for tremendous work organizing the questionnaire. This work was supported by the RADARS System, which is supported by subscriptions from pharmaceutical manufacturers and government and nongovernment agencies for surveillance, research, and reporting services. RADARS System is the property of Denver Health and Hospital Authority, a political subdivision of the State of Colorado. Denver Health retains exclusive ownership of all data, databases, and systems. No subscriber participated in the conception, analysis, drafting, or review of this manuscript. All authors contributed equally to the work.

Conflicts of Interest

Authors were employed by Denver Health and Hospital Authority during this work.

Multimedia Appendix 1
Supplemental methodological information and supporting results.
[PDF File (Adobe PDF File), 1150 KB - jmir_v21i10e15830_app1.pdf]

References

1. Reports and detailed tables from the 2017 National Survey on Drug Use and Health (NSDUH). Rockville: Substance Abuse and Mental Health Services Administration; 2018. URL: https://www.samhsa.gov/data/nsduh/reports-detailed-tables-2017-NSDUH [accessed 2019-09-29]
2. About the National Health Interview Survey. Hyattsville: National Center for Health Statistics URL: https://www.cdc.gov/nchs/nhis/about_nhis.htm [accessed 2019-09-29]
3. 2017 National Survey on Drug Use and Health: public use file codebook. Rockville: Substance Abuse and Mental Health Services Administration, Center for Behavioral Health Statistics and Quality URL: http://samhda.s3-us-gov-west-1.amazonaws.com/s3fs-public/field-uploads-protected/studies/NSDUH-2017/NSDUH-2017-datasets/NSDUH-2017-DS0001/NSDUH-2017-DS0001-info/NSDUH-2017-DS0001-info-codebook.pdf [accessed 2019-09-29]
4. Zipf G, Chiappa M, Porter KS, Ostchega Y, Lewis BG, Dostal J. National Health and Nutrition Examination Survey: plan and operations, 1999-2010. Vital Health Stat 1 2013 Aug(56):1-37 [FREE Full text] [Medline: 25078429]
5. Metzger DS, Koblin B, Turner C, Navaline H, Valenti F, Holte S, et al. Randomized controlled trial of audio computer-assisted self-interviewing: utility and acceptability in longitudinal studies. HIVNET Vaccine Preparedness Study Protocol Team. Am J Epidemiol 2000 Jul 15;152(2):99-106. [doi: 10.1093/aje/152.2.99] [Medline: 10909945]
6. Newman JC, Des Jarlais DC, Turner CF, Gribble J, Cooley P, Paone D. The differential effects of face-to-face and computer interview modes. Am J Public Health 2002 Feb;92(2):294-297. [Medline: 11818309]
7. van Gelder MMHJ, Bretveld RW, Roeleveld N. Web-based questionnaires: the future in epidemiology? Am J Epidemiol 2010 Dec 1;172(11):1292-1298 [FREE Full text] [doi: 10.1093/aje/kwq291] [Medline: 20880962]

8. Barber J, Kusunoki Y, Gatny H, Schulz P. Participation in an intensive longitudinal study with weekly web surveys over 2.5 years. J Med Internet Res 2016 Jun 23;18(6):e105 [FREE Full text] [doi: 10.2196/jmir.5422] [Medline: 27338859]
9. Erens B, Burkill S, Couper MP, Conrad F, Clifton S, Tanton C, et al. Nonprobability Web surveys to measure sexual behaviors and attitudes in the general population: a comparison with a probability sample interview survey. J Med Internet Res 2014 Dec 08;16(12):e276 [FREE Full text] [doi: 10.2196/jmir.3382] [Medline: 25488851]
10. Iachan R, Berman L, Kyle TM, Martin KJ, Deng Y, Moyse DN, et al. Weighting nonprobability and probability sample surveys in describing cancer catchment areas. Cancer Epidemiol Biomarkers Prev 2019 Mar;28(3):471-477. [doi: 10.1158/1055-9965.EPI-18-0797] [Medline: 30642842]
11. Miller EA, Berman L, Atienza A, Middleton D, Iachan R, Tortora R, et al. A feasibility study on using an internet-panel survey to measure perceptions of e-cigarettes in 3 metropolitan areas, 2015. Public Health Rep 2017;132(3):336-342 [FREE Full text] [doi: 10.1177/0033354917701888] [Medline: 28358990]
12. Anderson M, Perrin A, Jiang J, Kumar M. 10% of Americans don't use the internet. Who are they?. Washington: Pew Internet and American Life Project; 2019 Apr 22. URL: https://www.pewresearch.org/fact-tank/2019/04/22/some-americans-dont-use-the-internet-who-are-they/ [accessed 2019-09-29]
13. Curran PG. Methods for the detection of carelessly invalid responses in survey data. J Exp Soc Psychol 2016 Sep;66:4-19. [doi: 10.1016/j.jesp.2015.07.006]
14. Craig BM, Hays RD, Pickard AS, Cella D, Revicki DA, Reeve BB. Comparison of US panel vendors for online surveys. J Med Internet Res 2013 Nov 29;15(11):e260 [FREE Full text] [doi: 10.2196/jmir.2903] [Medline: 24292159]
15. Roster CA, Rogers RD, Albaum G, Klein D. A comparison of response characteristics from web and telephone surveys. Int J Market Res 2018 Jan 30;46(3):359-373. [doi: 10.1177/147078530404600301]
16. Leece P, Bhandari M, Sprague S, Swiontkowski MF, Schemitsch EH, Tornetta P, et al. Internet versus mailed questionnaires: a randomized comparison (2). J Med Internet Res 2004 Sep 24;6(3):e30 [FREE Full text] [doi: 10.2196/jmir.6.3.e30] [Medline: 15471756]
17. Yeager DS, Krosnick JA, Chang L, Javitz HS, Levendusky MS, Simpser A, et al. Comparing the accuracy of RDD telephone surveys and internet surveys conducted with probability and non-probability samples. Pub Opin Q 2011 Oct 05;75(4):709-747. [doi: 10.1093/poq/nfr020]
18. Kreuter F, McCulloch S, Presser S, Tourangeau R. The effects of asking filter questions in interleafed versus grouped format. Sociolog Methods Res 2011 Jan 11;40(1):88-104. [doi: 10.1177/0049124110392342]
19. Dillman D, Smyth J, Christian L. Internet, Phone, Mail, and Mixed-Mode Surveys: The Tailored Design Method. 4th Edition. Hoboken: John Wiley & Sons; 2014.
20. Meade AW, Craig SB. Identifying careless responses in survey data. Psychol Methods 2012 Sep;17(3):437-455. [doi: 10.1037/a0028085] [Medline: 22506584]
21. Deming WE, Stephan FF. On a least squares adjustment of a sampled frequency table when the expected marginal totals are known. Ann Math Statist 1940 Dec;11(4):427-444. [doi: 10.1214/aoms/1177731829]
22. Deville J, Sarndal C, Sautory O. Generalized raking procedures in survey sampling. J Am Statist Assoc 1993 Sep;88(423):1013. [doi: 10.2307/2290793]
23. Dart RC, Surratt HL, Cicero TJ, Parrino MW, Severtson SG, Bucher-Bartelson B, et al. Trends in opioid analgesic abuse and mortality in the United States. N Engl J Med 2015 Jan 15;372(3):241-248. [doi: 10.1056/NEJMsa1406143] [Medline: 25587948]
24. McCabe SE, Boyd CJ, Cranford JA, Morales M, Slayden J. A modified version of the Drug Abuse Screening Test among undergraduate students. J Subst Abuse Treat 2006 Oct;31(3):297-303 [FREE Full text] [doi: 10.1016/j.jsat.2006.04.010] [Medline: 16996392]
25. American Community Survey: 2014 Design and Methodology Report. Washington: US Census Bureau URL: http://www2.census.gov/programs-surveys/acs/methodology/design_and_methodology/acs_design_methodology_report_2014.pdf [accessed 2019-09-29]
26. Davenport E, El-Sanhurry N. Phi/Phimax: review and synthesis. Educ Psychol Meas 1991;51(4):1. [doi: 10.1177/0013164491051004003]
27. Dutwin D, Buskirk T. Apples to oranges or gala versus golden delicious? Pub Opin Q 2017;81(S1):213-239. [doi: 10.1093/poq/nfw061]
28. Kennedy C, Mercer A, Keeter S, Hatley N, McGeeney K, Gimenez A. Evaluating online nonprobability surveys. Washington: Pew Internet and American Life Project; 2016 May 02. URL: https://www.pewresearch.org/wp-content/uploads/2016/04/Nonprobability-report-May-2016-FINAL.pdf [accessed 2019-09-29]
29. Izrael D, Hoaglin D, Battaglia M. A SAS macro for balancing a weighted sample. 2000. URL: https://support.sas.com/resources/papers/proceedings/proceedings/sugi25/25/st/25p258.pdf [accessed 2019-09-29]
30. Lohr S. Sampling: Design and Analysis. 2nd Edition. Boston: Cengage; 2009.
31. Andreeva VA, Deschamps V, Salanave B, Castetbon K, Verdot C, Kesse-Guyot E, et al. Comparison of dietary intakes between a large online cohort study (Etude NutriNet-Santé) and a nationally representative cross-sectional study (Etude Nationale Nutrition Santé) in France: addressing the issue of generalizability in e-epidemiology. Am J Epidemiol 2016 Nov 01;184(9):660-669 [FREE Full text] [doi: 10.1093/aje/kww016] [Medline: 27744386]

32. De Vera MA, Ratzlaff C, Doerfling P, Kopec J. Reliability and validity of an internet-based questionnaire measuring lifetime physical activity. Am J Epidemiol 2010 Nov 15;172(10):1190-1198 [FREE Full text] [doi: 10.1093/aje/kwq273] [Medline: 20876666]
33. Russell CW, Boggs DA, Palmer JR, Rosenberg L. Use of a web-based questionnaire in the Black Women's Health Study. Am J Epidemiol 2010 Dec 1;172(11):1286-1291 [FREE Full text] [doi: 10.1093/aje/kwq310] [Medline: 20937635]
34. Greenacre ZA. The importance of selection bias in internet surveys. Open J Statist 2016;06(03):397-404. [doi: 10.4236/ojs.2016.63035]
35. Tourangeau R, Conrad FG, Couper MP. The Science of Web Surveys. Oxford: Oxford University press; 2013.
36. Elliott MR, Valliant R. Inference for nonprobability samples. Statist Sci 2017 May;32(2):249-264. [doi: 10.1214/16-STS598]
37. Schonlau M. Will web surveys ever become part of mainstream research? J Med Internet Res 2004 Sep 23;6(3):e31 [FREE Full text] [doi: 10.2196/jmir.6.3.e31] [Medline: 15471757]
38. Dasgupta N, Beletsky L, Ciccarone D. Opioid crisis: no easy fix to its social and economic determinants. Am J Public Health 2018 Feb;108(2):182-186. [doi: 10.2105/AJPH.2017.304187] [Medline: 29267060]
39. Olsen EO, O'Donnell J, Mattson CL, Schier JG, Wilson N. Notes from the field: unintentional drug overdose deaths with kratom detected—27 states, July 2016–December 2017. MMWR Morb Mortal Wkly Rep 2019 Apr 12;68(14):326-327 [FREE Full text] [doi: 10.15585/mmwr.mm6814a2] [Medline: 30973850]
40. El Zahran T, Schier J, Glidden E, Kieszak S, Law R, Bottei E, et al. Characteristics of tianeptine exposures reported to the National Poison Data System—United States, 2000-2017. MMWR Morb Mortal Wkly Rep 2018 Aug 03;67(30):815-818 [FREE Full text] [doi: 10.15585/mmwr.mm6730a2] [Medline: 30070980]
41. Kariisa M, Scholl L, Wilson N, Seth P, Hoots B. Drug overdose deaths involving cocaine and psychostimulants with abuse potential—United States, 2003-2017. MMWR Morb Mortal Wkly Rep 2019 May 03;68(17):388-395 [FREE Full text] [doi: 10.15585/mmwr.mm6817a3] [Medline: 31048676]
42. Evans EA, Sullivan MA. Abuse and misuse of antidepressants. Subst Abuse Rehabil 2014;5:107-120 [FREE Full text] [doi: 10.2147/SAR.S37917] [Medline: 25187753]
43. Hawkins KL, Gidal BE. When adverse effects are seen as desirable: abuse potential of the newer generation antiepileptic drugs. Epilepsy Behav 2017 Dec;77:62-72. [doi: 10.1016/j.yebeh.2017.10.007] [Medline: 29111505]

Abbreviations

ACS: American Community Survey
CHERRIES: Checklist for Reporting Results of Internet E-Surveys
DAST-10: 10-item Drug Abuse Screening Test
NHIS: National Health Interview Survey
NMU: nonmedical use
NMURx: Survey of Non-Medical Use of Prescription Drugs
NSDUH: National Survey of Drug Use and Health
RADARS: Researched Abuse, Diversion and Addiction-Related Surveillance

Edited by G Eysenbach; submitted 13.08.19; peer-reviewed by X Chen, A Dregan; comments to author 05.09.19; revised version received 11.09.19; accepted 18.09.19; published 25.10.19.

Please cite as:
Black JC, Rockhill K, Forber A, Amioka E, May KP, Haynes CM, Dasgupta N, Dart RC
An Online Survey for Pharmacoepidemiological Investigation (Survey of Non-Medical Use of Prescription Drugs Program): Validation Study
J Med Internet Res 2019;21(10):e15830
URL: http://www.jmir.org/2019/10/e15830/
doi:10.2196/15830
PMID:31654568

Original Paper

Characterizing the Rural Opioid Use Environment in Kentucky Using Google Earth: Virtual Audit

Natalie Danielle Crawford[1], PhD; Regine Haardöerfer[1], PhD; Hannah Cooper[1], PhD; Izraelle McKinnon[2], MPH; Carla Jones-Harrell[1], MPhil; April Ballard[3], BA; Sierra Shantel von Hellens[4], BA; April Young[3], PhD

[1]Department of Behavioral Sciences and Health Education, Rollins School of Public Health, Emory University, Atlanta, GA, United States

[2]Department of Epidemiology, Emory University, Atlanta, GA, United States

[3]Department of Environmental Health, Rollins School of Public Health, Emory University, Atlanta, GA, United States

[4]Morehead State University, Morehead, KY, United States

Corresponding Author:
Natalie Danielle Crawford, PhD
Department of Behavioral Sciences and Health Education, Rollins School of Public Health, Emory University
1518 Clifton Road
Atlanta, GA
United States
Phone: 1 4047429445
Email: ndcrawford@emory.edu

Abstract

Background: The opioid epidemic has ravaged rural communities in the United States. Despite extensive literature relating the physical environment to substance use in urban areas, little is known about the role of physical environment on the opioid epidemic in rural areas.

Objective: This study aimed to examine the reliability of Google Earth to collect data on the physical environment related to substance use in rural areas.

Methods: Systematic virtual audits were performed in 5 rural Kentucky counties using Google Earth between 2017 and 2018 to capture land use, health care facilities, entertainment venues, and businesses. In-person audits were performed for a subset of the census blocks.

Results: We captured 533 features, most of which were images taken before 2015 (71.8%, 383/533). Reliability between the virtual audits and the gold standard was high for health care facilities (>83%), entertainment venues (>95%), and businesses (>61%) but was poor for land use features (>18%). Reliability between the virtual audit and in-person audit was high for health care facilities (83%) and entertainment venues (62%) but was poor for land use (0%) and businesses (12.5%).

Conclusions: Poor reliability for land use features may reflect difficulty characterizing features that require judgment or natural changes in the environment that are not reflective of the Google Earth imagery because it was captured several years before the audit was performed. Virtual Google Earth audits were an efficient way to collect rich neighborhood data that are generally not available from other sources. However, these audits should use caution when the images in the observation area are dated.

(J Med Internet Res 2019;21(10):e14923) doi:10.2196/14923

KEYWORDS

opioid-related disorders; rural health; built environment

Introduction

The opioid epidemic has had a devastating impact on Americans, resulting in increased levels of addiction and overdose, particularly among those who live in rural areas [1,2]. Although drug use has been historically perceived as an urban problem, the epidemiology of substance use and drug overdose has shifted substantially from cities to rural areas [3,4]. For example, in Kentucky alone, a state that is about 71% rural [5], synthetic opioid and heroin overdose deaths have increased to 780 deaths and 269 deaths in 2017, representing a 10-fold and 2-fold increase since 2013, respectively [6]. Moreover, the rate of drug overdose in rural areas has surpassed that of many urban areas [7,8], with Kentucky having the fourth highest overdose rate nationally [9].

Rural areas cover approximately 97% of the land area in the United States [10], and despite the salience of the physical environment for substance use and related harms in urban areas, its impact in rural areas has received significantly less attention by the scientific community [11]. There is an extensive body of literature in urban environments that has consistently shown how substance use, injection-risk behavior, and HIV transmission are related to the physical environment, which includes both built and natural elements [12-19]. This research is supported by broken windows and risk environment theories that describe how visible decay such as run-down housing in a neighborhood results in crime and disorder [20], which is consonant with the contextual environment where these risks occur [21]. In urban areas, researchers have relied on existing administrative geographical data [22], in-person audits, and, more recently, virtual audits from public data sources with video or satellite imagery to assess the physical environment [23-25].

Analysis of existing data from administrative sources such as the US Census allows researchers to bypass the efforts involved in primary data collection; however, these secondary data often fail to capture constructs that are critical for understanding many determinants of health. This is particularly true in rural and international settings where there is a unique context in which health and health behaviors are produced [26]. For example, our qualitative data suggest that self-service drive-through car washes, which are often a part of gas stations, are used as a private area for injection drug use in rural areas. Although existing business data might indicate that a gas station exists in an area, refined data on the presence of a car wash within that gas station are rarely available. In addition, administrative data sources often create measures for geographic units based on population size. Therefore, in rural areas where the population density is small, the geographic unit covers an expansive area that may lack precision and utility. Measuring characteristics of rural areas may thus require that we develop novel measures and innovative data collection strategies to fully capture the context in rural areas for the appropriate geographic exposure area.

In-person audits are commonly used when existing data sources fail to capture specific characteristics and tend to be the strongest at assessing features of the physical environment [22,27]. However, they require trained researchers who directly observe and document specific characteristics of the physical environment, and this process can be very time-consuming and costly [28,22], particularly for rural areas, which generally have a large landmass that requires long travel times [26,29]. Importantly, in-person audits in rural areas may also be limited because of poorly vascularized roads that make it difficult to identify features of the environment that are not directly off a driving path. In urban areas, there are often enough streets to connect areas that overcome this challenge.

Virtual audits using Google Earth have shown great promise in overcoming the inefficiency of in-person audits [22,26,30,31]. Google Earth is a free 3-dimensional geographic program that provides aerial, satellite, and street view imagery on the Web and covers a vast proportion of the world's surface area [32]. Virtual *walk arounds* of an area in Google Earth have been performed systematically to capture specific characteristics of the physical environment. Recent studies using Google Earth audits have shown strong interrater reliability and concurrent validity for a number of characteristics in both national and international settings [22,33], including food environment, recreational facilities, and street characteristics. Yet, most of these studies have been conducted in urban environments, and the utility of this innovative tool for examining rural environments is unclear. In rural areas, virtual audits using Google Earth have provided measures that are comparable with in-person audits for a number of characteristics including the presence of sidewalks [34] and the number of objective housing characteristics related to healthy aging [35]. To our knowledge, no studies have examined how well Google Earth audits perform when examining characteristics that promote or reduce risks related to substance use in rural environments.

The purpose of this study was to describe the utility and reliability of performing virtual audits using Google Earth technology to measure features of the physical environment that might be related to nonmedical prescription opioid use, heroin use, injection drug use, health care use, and overdose in a rural area. We assessed interrater reliability of built environment audits using Google Earth technology and interrater reliability of Google Earth audits compared with in-person audits [36]. We use this information to make recommendations for the use of Google Earth technology in neighborhood data collection specific to rural areas.

Methods

Overview

Neighborhood audits via Google Earth and in-person took place within a larger study that examined the influence of the risk environment on opioid use among young adults (aged 18-35 years) in 5 counties in rural Kentucky through a partnership between the University of Kentucky and Emory University. These counties were chosen in the parent study because of high levels of nonmedical prescription opioid use, overdose, and poverty. Neighborhood audits were conducted between July 2017 and August 2018. Herein, we describe the virtual audit training, virtual audit data collection, virtual audit reconciliation, and the in-person audit.

Virtual Audit Training

Auditors underwent extensive training before data collection. All auditors were advanced graduate students who had taken coursework in spatial data. Auditors were required to review a step-by-step protocol of the audit methods. An in-depth discussion of the definition and appropriate classification of each neighborhood characteristic was reviewed with each auditor. Independent and supervised practice audits of sample areas were then performed. Weekly discussion and troubleshooting of the data collection were performed. Additional training of the auditors was performed as needed.

Virtual Audit Data Collection

Systematic auditing of 49 census blocks was performed for 5 rural Kentucky counties using cartographic boundary files from the 2016 US Census [37]. Census blocks are small geographical areas nested within census tracts, but larger than city blocks,

for which basic demographic data can be obtained for a population. Similar to in-person neighborhood audits, Google Street View audits were systematically initiated at the same location (eg, most southeastern point) for each county to capture physical environment characteristics that might be related to nonmedical prescription opioid use, heroin use, injection drug use, health care use, and overdose among young adults living in rural areas. Previous studies have examined each block face of an area to collect the desired measures [22]. Street view images were supplemented with aerial images in Google Earth for select characteristics.

To inform the functionality of Google Earth for virtual audits in rural areas, we also collected data on the date on which the images were captured for each block to understand how well these images matched with the present time frame and in-person audits. Auditors also maintained a log of the time required to audit each block. The auditors noted when they did not know whether a location fell within a physical environment category, experienced a visual problem with Google Earth (ie, resolution issue), or if a characteristic was visible in 1 view (aerial or street view) but not in another. Each block was audited by 2 independent, trained auditors.

Virtual Audit Reconciliation

We established a final virtual audit dataset after reconciling discrepancies between the 2 independent Google Earth auditors. Discrepancies were identified by comparing the quantity of each physical environment characteristic across the 2 independent auditors for each block. Then, a third independent Google Earth auditor, who was considered the gold standard, repeated the audit of the entire block to specifically identify the correct number of physical environment characteristics for which discrepancies were present. The third independent auditor was also able to compare the specific latitude and longitude, where the 2 independent auditors disagreed. Disagreements and reconciliation findings were reviewed daily between the third independent auditor and the senior author to finalize all discrepancies.

In-Person Audits

To validate the data obtained from the virtual audits, in-person audits were conducted for a subsample of the census blocks where virtual audits were performed. The subsample was selected by excluding census blocks with Google Street View images that were only collected before 2015 to ensure that the neighborhood characteristics were a good representation of the current environment, given that some features of the environment can disappear and appear. Then, we further excluded the sample based on counties with at least 800 residents to ensure that there was an adequate number of physical environment characteristics for comparison with the virtual audits. A total of 8 blocks were randomly selected across all 5 counties.

We hired a resident of 1 of the counties to perform in-person audits. We provided extensive training to the in-person auditor to ensure systematic and standardized collection of each characteristic similar to the virtual audit. Driving routes were created with physical maps of each block starting at the most southeastern point of each block. A data collection tool was developed using the Fulcrum app [38], which was installed on a mobile phone to capture the latitude and longitude of each neighborhood characteristic. The Fulcrum app is ideal for rural settings where internet connectivity may be limited as it can collect data without a constant internet connection and can automatically produce time stamps of the data collection session [39].

Physical Environment Characteristics Assessed in Virtual and In-Person Audits

Data were collected on various characteristics of the physical environment and categorized based on whether they represented land use, a health care facility, entertainment venue, or business. The specific features captured within each of these categories are shown in Textbox 1. In brief, land use characteristics are made up of features that describe the function of property, health care facilities represent buildings where an individual can obtain health care, entertainment venues represent businesses that individuals patron for pleasure, and businesses represent storefronts where individuals obtain day-to-day necessities. Importantly, categorization of each feature was based on the literature and our qualitative data that informed data collection of characteristics that were unique to the Kentucky area (eg, hollows). Existing literature has shown that land use, entertainment venues, and businesses (eg, those that allow open alcohol use or provide opportunities for sex) may promote opportunities to obtain and use illicit substances [17,40,41], whereas health care facilities may be protective and provide opportunities to reduce substance use and injection risk behaviors [13]. Given the small numbers of each feature, we combined them across each category for presentation. Disaggregated data that are deidentified for each county are available on request.

Textbox 1. Physical environment features assessed in virtual and in-person audits in each category.

Land use

- Boarded up and dilapidated business [42]
- Boarded up and dilapidated homes [22,42]
- Defunct mines and industrial sites (captured using aerial views)
- Water recreation areas (captured using aerial views) [22,42]
- Hollows (captured using aerial views)
- Trailer parks (captured using aerial views)

Health care facilities

- Drug-related and HIV-related health care sites
- General health care sites [25]
- Pharmacies [22]
- Syringe/needle exchange programs

Entertainment venues

- Liquor store [22,43]
- Tobacco store
- Motel/hotel

Businesses

- Fast food restaurants [22,44]
- General restaurants [44]
- Gas stations [22,44]
- Car washes
- Cemeteries (captured using aerial views)
- Truck stops

Data Analysis

First, we calculated the frequencies of the images collected each year. Audit time from the virtual audits was summed for each block and averaged across all blocks in each county. We describe the median and interquartile range (IQR) for each physical environment category. Owing to the sparsity of the features measured in these rural communities and the small number of census blocks, traditional intraclass correlations (ICCs) were unstable and vastly underestimated between-auditor agreement [45,46]. Therefore, we calculated the percent agreement of each audit based on complete agreement between the 2 independent auditors.

We also calculated the percent at which each auditor overestimated or underestimated each feature based on the virtual gold standard. If an auditor identified the same number of a feature within a given block as the gold standard, this auditor's scoring was coded a *match*. If the auditor identified a higher or lower number of a feature than the gold standard, the auditor's scoring was coded *over* and *under*, respectively. The percentages were calculated as number of segments for which an auditor identified *over*, *under*, or *matched* with the gold standard. Percent agreement was also reconciled between the final virtual audit data, compared with the gold standard auditor, and the in-person audits.

Results

Textbox 1 describes each feature that was collected for each physical environment category in the virtual and in-person audits. A total of 533 data points were captured in the virtual audits, 383 (72%) of which were from images taken before 2015 with more than 40% of those dating back to 2009 (data not shown). Table 1 shows the median and IQR for each characteristic captured in the final (reconciled) virtual audit data. In general, the data for these rural areas are sparse. There was a median of 14 (IQR 9-25) features in each block in the land use category. Boarded up and dilapidated homes (median 8, IQR 3-8) were the most frequently observed feature in the land use category and overall. There was a median of 0 (IQR 0-1) health care facilities per block, where each feature was equally underrepresented in each block. There was a median of 1 (IQR 1-3) entertainment venue where each feature within this category was also equally sparse. There was a median of 6 businesses (IQR 2-9), which was mostly represented by cemeteries (median 2; IQR 1-2). There were no syringe/needle exchange programs identified in any county, and liquor stores

were only present in 1 county, which was the only county where alcohol could be sold in our geographic sample.

The median number of Google Earth images that were difficult to decipher because of poor resolution of the images or missing street view images where aerial images were available was also small, but the disaggregated data showed that missingness was more common in the least populated areas. The average audit time per block was 1.97 hours, ranging from 58 min to 227 min per block. In-person audits required about 8 hours of auditing time per block, not including the time required to map each area.

Table 2 shows the reliability of the physical environment characteristics for the virtual and in-person audits. The reconciled data comparing the 2 independent virtual audits with the gold standard show that agreement was high for health care facilities (>83%), entertainment venues (>95%), and businesses (>61%). Agreement was poor for land use features (>18%) and varied substantially between auditors (36.7% vs 18.4%). Land use features included boarded up and dilapidated homes, businesses and defunct mines and industrial sites, and hollows. The independent auditors tended to underestimate (range: 2%-73.5%) versus overestimate (range: 0%-8.2%) the presence of a feature compared with the gold standard. For features in the land use category, these overestimates were high (range: 61.2%-73.5%), whereas health care facilities (range: 10.6%-16.7%), entertainment venues (range: 2%-4.1%), and businesses (range: 18.4%-34.7%) had lower estimates.

When comparing the reconciled virtual audit data with the in-person audit data, reliability was high for health care facilities (83%) and entertainment venues (62%) but was poor for land use (0%) and businesses (12.5%). The in-person audit data generally underestimated features in the land use category (75%) and health care facilities (16.7%) compared with the reconciled virtual data. However, features in the entertainment venue and business categories were overestimated (25% and 62.5%, respectively) compared with the reconciled virtual data.

Table 1. Descriptive neighborhood characteristics collected using virtual audits for 5 counties in Kentucky (N=64,061).

Characteristics	All counties (n=64)
Land use, mean (IQR[a])	14 (9-25)
Health care facilities, mean (IQR)	0 (0-1)
Entertainment venues, mean (IQR)	1 (1-3)
Businesses, mean (IQR)	6 (2-9)
Do not know, mean (IQR)	1 (0-4)
Unclear, mean (IQR)	0 (0-0)
Missing images, mean (IQR)	2 (0-3)
Virtual audit time per block (hours), mean (SD)	1.97 (1.07)
In-person audit time per block (hours), mean (SD)	8 (1.92)

[a]IQR: interquartile range.

Table 2. Reliability of neighborhood characteristics collected in rural Kentucky using virtual and in-person audits versus the gold standard.

Characteristics	Virtual audits		In-person audits, % agreement
	Audit 1, % agreement	Audit 2, % agreement	
Land use			
Over	2.0	8.2	25.0
Match	36.7	18.4	0.0
Under	61.2	73.5	75.0
Health care facilities			
Over	0.0	0.0	0.0
Match	89.4	83.3	83.0
Under	10.6	16.7	16.7
Entertainment venue			
Over	2.0	0.0	25.0
Match	95.9	95.9	62.5
Under	2.0	4.1	12.5
Businesses			
Over	6.1	4.1	62.5
Match	75.5	61.2	12.5
Under	18.4	34.7	25.0

Discussion

Principal Findings

Virtual audits using Google Earth show strong potential for assessing the built environment in rural areas for objective features such as health care facilities, entertainment venues, and businesses. However, virtual audits are less reliable when the characteristic requires some judgment or may change within a short period such as dilapidated housing or features that are no longer in operation. Other studies have similarly encountered difficulties when assessing characteristics of the built environment that are more ambiguous, such as the amount of loitering on the street or graffiti [26,47]. In our study, virtual audit reliability was the lowest for the land use category, which included boarded up and dilapidated homes and businesses. This may reflect the difficulty in capturing subjective features that require some judgment to assess. However, the aerial imagery in Google did allow for better view of mines, which would have been difficult, if not impossible, to assess in street view. Thus, the flexibility of street and aerial imagery is a strength.

Challenges and Limitations

Our audits in rural environments revealed some unique challenges that should be considered in future virtual audits. First, older and missing images in Google Earth are an important limitation. Google states that street view audits are performed every 2 to 3 years to maintain updated imaging [33]. However, a substantial proportion (72%) of the images we analyzed were taken before 2015 and dated as far back as 2008, 9-10 years before the virtual audit was done. In these instances, specific issues were that the street view data were unavailable or incomplete, and for some characteristics, such as strip malls, they were also unavailable in the aerial view. It is unclear if Google Earth fails to update their images in places with smaller populations or other characteristics, but we found older images in all of the counties assessed in our virtual audit regardless of how populated the county was. Older images made it difficult to estimate reliability between our virtual and in-person audits. For example, lower agreement may reflect images that were not captured by an auditor because of auditor failure, or it could reflect that a characteristic appeared or disappeared in the years since the Google Earth image was taken. To attempt to reduce this bias, we performed in-person audits in census blocks with images captured post 2015. However, it is still possible that homes or businesses can become run-down or be torn down completely within just a couple of years [48]. In line with this hypothesis, the in-person auditor tended to report an underestimate of features in the land use category and overestimate of features in the business category. Of note, our audits could only identify features that were visible through external signage of the feature. In many areas, health services such as syringe exchange are often not advertised widely to avoid negative attention. Therefore, syringe exchange programs that were located in health departments were not captured through our audits. It is important to note that this would also be a limitation of in-person audits.

By their very nature, rural environments will produce fewer observations than urban environments for many features that research shows are relevant to substance misuse and related harms. However, we must note that as research on physical environment and drug use expands, new influential features of rural areas may be identified that are more commonplace in these areas. The sparsity of the features assessed here poses

difficulty when calculating standard reliability measures. ICCs severely underestimate agreement when the data are sparse [45,46]. Therefore, we were only able to quantify percent agreement, but these estimates may also be affected by the small sample size.

In addition, Google Earth frequently crashed, which resulted in loss of work. Furthermore, auditor fatigue may be particularly salient in rural audits because rural areas have fewer features to assess, so the audits may be more cumbersome. Rural areas also generally have a larger landmass than urban areas and require more time to thoroughly review than an urban area. This may contribute to auditor turn over, resulting in the need for additional training and resources for new auditors.

Nevertheless, virtual audits are an efficient way to collect rich neighborhood data that are not routinely available through other sources. Google Earth is particularly viable in rural areas where there is a need to capture features that are distant from roads that are less connected than urban areas and would be difficult if not impossible to reach via in-person audits. Thus, Google Earth is a powerful tool that should be considered for future research and data collection in rural environments. These data will allow for research that elucidates key features of rural environments related to opioid use and points to critical points of intervention to reduce substance misuse and related harms.

Acknowledgments

The authors would like to acknowledge their funding sources: NIDA R21 DA042727-01 and the Emory University Prevention Science Core Center for AIDS Research (P30AI050409). They would also like to thank their virtual and in-person auditors for their time and dedication to this study.

Conflicts of Interest

None declared.

References

1. Kaiser Health News. 2016. Surgeon General Murthy Wants America To Face Up To Addiction URL: https://khn.org/news/surgeon-general-murthy-wants-america-to-face-up-to-addiction/ [accessed 2017-01-01]
2. Noonan R. Centers for Disease Control and Prevention. 2017. Rural America in Crisis: The Changing Opioid Overdose Epidemic URL: https://blogs.cdc.gov/publichealthmatters/2017/11/opioids/ [accessed 2017-01-01]
3. Havens JR, Oser CB, Leukefeld CG. Increasing prevalence of prescription opiate misuse over time among rural probationers. J Opioid Manag 2007;3(2):107-111. [doi: 10.5055/jom.2007.0047] [Medline: 17520990]
4. Paulozzi LJ, Xi Y. Recent changes in drug poisoning mortality in the United States by urban-rural status and by drug type. Pharmacoepidemiol Drug Saf 2008 Oct;17(10):997-1005. [doi: 10.1002/pds.1626] [Medline: 18512264]
5. Davis AF. The Community and Economic Development Initiative of Kentucky. 2009. Kentucky's Urban/Rural Landscape: What is Driving the Differences in Wealth Across Kentucky? URL: https://cedik.ca.uky.edu/files/kyruralurbanwealthdifferences.pdf [accessed 2019-08-21]
6. National Institute on Drug Abuse (NIDA). 2018. Kentucky Opioid Summary URL: https://www.drugabuse.gov/opioid-summaries-by-state/kentucky-opioid-summary [accessed 2017-12-01]
7. Keyes KM, Cerdá M, Brady JE, Havens JR, Galea S. Understanding the rural-urban differences in nonmedical prescription opioid use and abuse in the United States. Am J Public Health 2014 Feb;104(2):e52-e59. [doi: 10.2105/AJPH.2013.301709] [Medline: 24328642]
8. Mosher H, Zhou Y, Thurman AL, Sarrazin MV, Ohl ME. Trends in hospitalization for opioid overdose among rural compared to urban residents of the United States, 2007-2014. J Hosp Med 2017 Nov;12(11):925-929. [doi: 10.12788/jhm.2793] [Medline: 29091981]
9. Centers for Disease Control and Prevention. 2019. Drug Overdose Mortality by State URL: https://www.cdc.gov/nchs/pressroom/sosmap/drug_poisoning_mortality/drug_poisoning.htm [accessed 2018-01-01]
10. Charreire H, Mackenbach JD, Ouasti M, Lakerveld J, Compernolle S, Ben-Rebah M, et al. Using remote sensing to define environmental characteristics related to physical activity and dietary behaviours: a systematic review (the SPOTLIGHT project). Health Place 2014 Jan;25:1-9. [doi: 10.1016/j.healthplace.2013.09.017] [Medline: 24211730]
11. Anguelov D, Dulong C, Filip D, Frueh C, Lafon S, Lyon R, et al. Google street view: capturing the world at street level. Computer 2010 Jun;43(6):32-38. [doi: 10.1109/MC.2010.170]
12. Mason MJ, Korpela K. Activity spaces and urban adolescent substance use and emotional health. J Adolesc 2009 Aug;32(4):925-939. [doi: 10.1016/j.adolescence.2008.08.004] [Medline: 18851877]
13. Linas BS, Latkin C, Westergaard RP, Chang LW, Bollinger RC, Genz A, et al. Capturing illicit drug use where and when it happens: an ecological momentary assessment of the social, physical and activity environment of using versus craving illicit drugs. Addiction 2015 Feb;110(2):315-325 [FREE Full text] [doi: 10.1111/add.12768] [Medline: 25311241]

14. Cerdá M, Ransome Y, Keyes KM, Koenen KC, Tardiff K, Vlahov D, et al. Revisiting the role of the urban environment in substance use: the case of analgesic overdose fatalities. Am J Public Health 2013 Dec;103(12):2252-2260 [FREE Full text] [doi: 10.2105/AJPH.2013.301347] [Medline: 24134362]
15. Roberts ET, Friedman SR, Brady JE, Pouget ER, Tempalski B, Galea S. Environmental conditions, political economy, and rates of injection drug use in large US metropolitan areas 1992-2002. Drug Alcohol Depend 2010 Jan 15;106(2-3):142-153 [FREE Full text] [doi: 10.1016/j.drugalcdep.2009.08.007] [Medline: 19748745]
16. Bernstein KT, Galea S, Ahern J, Tracy M, Vlahov D. The built environment and alcohol consumption in urban neighborhoods. Drug Alcohol Depend 2007 Dec 1;91(2-3):244-252. [doi: 10.1016/j.drugalcdep.2007.06.006] [Medline: 17644274]
17. Galea S, Rudenstine S, Vlahov D. Drug use, misuse, and the urban environment. Drug Alcohol Rev 2005 Mar;24(2):127-136. [doi: 10.1080/09595230500102509] [Medline: 16076582]
18. Hembree C, Galea S, Ahern J, Tracy M, Piper TM, Miller J, et al. The urban built environment and overdose mortality in New York city neighborhoods. Health Place 2005 Jun;11(2):147-156. [doi: 10.1016/j.healthplace.2004.02.005] [Medline: 15629682]
19. Friedman SR, Mateu-Gelabert P, Ruggles KV, Goodbody E, Syckes C, Jessell L, et al. Sexual risk and transmission behaviors, partnerships and settings among young adult nonmedical opioid users in New York city. AIDS Behav 2017 Apr;21(4):994-1003 [FREE Full text] [doi: 10.1007/s10461-016-1672-7] [Medline: 28058567]
20. Kelling GL, Wilson JQ. The Atlantic. 1982. Broken Windows: The Police and Neighborhood Safety URL: https://www.theatlantic.com/magazine/archive/1982/03/broken-windows/304465/ [accessed 2014-01-01]
21. Rhodes T. Risk environments and drug harms: a social science for harm reduction approach. Int J Drug Policy 2009 May;20(3):193-201. [doi: 10.1016/j.drugpo.2008.10.003] [Medline: 19147339]
22. Clarke P, Ailshire J, Melendez R, Bader M, Morenoff J. Using Google Earth to conduct a neighborhood audit: reliability of a virtual audit instrument. Health Place 2010 Nov;16(6):1224-1229 [FREE Full text] [doi: 10.1016/j.healthplace.2010.08.007] [Medline: 20797897]
23. Marco M, Gracia E, Martín-Fernández M, López-Quílez A. Validation of a Google street view-based neighborhood disorder observational scale. J Urban Health 2017 Apr;94(2):190-198 [FREE Full text] [doi: 10.1007/s11524-017-0134-5] [Medline: 28236183]
24. He L, Páez A, Liu D. Built environment and violent crime: an environmental audit approach using Google Street view. Comput Environ Urban 2017 Nov;66:83-95. [doi: 10.1016/j.compenvurbsys.2017.08.001]
25. Pliakas T, Hawkesworth S, Silverwood RJ, Nanchahal K, Grundy C, Armstrong B, et al. Optimising measurement of health-related characteristics of the built environment: comparing data collected by foot-based street audits, virtual street audits and routine secondary data sources. Health Place 2017 Jan;43:75-84 [FREE Full text] [doi: 10.1016/j.healthplace.2016.10.001] [Medline: 27902960]
26. Bader MD, Mooney SJ, Lee YJ, Sheehan D, Neckerman KM, Rundle AG, et al. Development and deployment of the computer assisted neighborhood visual assessment system (CANVAS) to measure health-related neighborhood conditions. Health Place 2015 Jan;31:163-172 [FREE Full text] [doi: 10.1016/j.healthplace.2014.10.012] [Medline: 25545769]
27. Schaefer-McDaniel N, Caughy MO, O'Campo P, Gearey W. Examining methodological details of neighbourhood observations and the relationship to health: a literature review. Soc Sci Med 2010 Jan;70(2):277-292. [doi: 10.1016/j.socscimed.2009.10.018] [Medline: 19883966]
28. Mooney SJ, Bader MD, Lovasi GS, Teitler JO, Koenen KC, Aiello AE, et al. Street audits to measure neighborhood disorder: virtual or in-person? Am J Epidemiol 2017 Aug 1;186(3):265-273 [FREE Full text] [doi: 10.1093/aje/kwx004] [Medline: 28899028]
29. Brownson RC, Hoehner CM, Day K, Forsyth A, Sallis JF. Measuring the built environment for physical activity: state of the science. Am J Prev Med 2009 Apr;36(4 Suppl):S99-123.e12 [FREE Full text] [doi: 10.1016/j.amepre.2009.01.005] [Medline: 19285216]
30. Aghaabbasi M, Moeinaddini M, Shah MZ, Asadi-Shekari Z. Addressing issues in the use of Google tools for assessing pedestrian built environments. J Transp Geogr 2018 Dec;73:185-198. [doi: 10.1016/j.jtrangeo.2018.10.004]
31. Wilson JS, Kelly CM, Schootman M, Baker EA, Banerjee A, Clennin M, et al. Assessing the built environment using omnidirectional imagery. Am J Prev Med 2012 Feb;42(2):193-199 [FREE Full text] [doi: 10.1016/j.amepre.2011.09.029] [Medline: 22261217]
32. Google. 2014. Google Earth URL: https://www.google.com/earth/learn/beginner.html [accessed 2013-01-01]
33. Curtis JW, Curtis A, Mapes J, Szell AB, Cinderich A. Using Google Street view for systematic observation of the built environment: analysis of spatio-temporal instability of imagery dates. Int J Health Geogr 2013 Dec 3;12:53 [FREE Full text] [doi: 10.1186/1476-072X-12-53] [Medline: 24298903]
34. Janssen I, Rosu A. Measuring sidewalk distances using Google Earth. BMC Med Res Methodol 2012 Mar 29;12:39 [FREE Full text] [doi: 10.1186/1471-2288-12-39] [Medline: 22458340]
35. Burton EJ, Mitchell L, Stride CB. Good places for ageing in place: development of objective built environment measures for investigating links with older people's wellbeing. BMC Public Health 2011 Nov 1;11:839 [FREE Full text] [doi: 10.1186/1471-2458-11-839] [Medline: 22044518]

36. Hallgren K. Computing inter-rater reliability for observational data: an overview and tutorial. Tutor Quant Methods Psychol 2012;8(1):23-34 [FREE Full text] [doi: 10.20982/tqmp.08.1.p023] [Medline: 22833776]
37. United States Census Bureau. 2010. Cartographic Boundary Files - Shapefile URL: https://www.census.gov/geographies/mapping-files/time-series/geo/carto-boundary-file.html [accessed 2015-01-01]
38. Fulcrum. 2018. Mobile Location Leverage URL: https://web.fulcrumapp.com/shares/0cf771b8af8b7b12/videos/2d6d3e4c-9d00-4e6f-a7a4-87fa2a0cf606/play [accessed 2018-01-01]
39. Wilson P. Fulcrum. 2012. Offline Imagery in Fulcrum URL: http://www.fulcrumapp.com [accessed 2018-01-01]
40. Mason MJ, Mennis J, Coatsworth JD, Valente TW, Lawrence F, Pate P. The relationship of place to substance use and perceptions of risk and safety in urban adolescents. J Environ Psychol 2009 Dec 1;29(4):485-492 [FREE Full text] [doi: 10.1016/j.jenvp.2009.08.004] [Medline: 20161426]
41. Mennis J, Stahler GJ, Mason MJ. Risky substance use environments and addiction: a new frontier for environmental justice research. Int J Environ Res Public Health 2016 Jun 18;13(6):E607 [FREE Full text] [doi: 10.3390/ijerph13060607] [Medline: 27322303]
42. Rundle AG, Bader MD, Richards CA, Neckerman KM, Teitler JO. Using Google Street view to audit neighborhood environments. Am J Prev Med 2011 Jan;40(1):94-100 [FREE Full text] [doi: 10.1016/j.amepre.2010.09.034] [Medline: 21146773]
43. Bowers Y, Rendall-Mkosi K, Davids A, Nel E, Jacobs N, London L. Liquor outlet density, deprivation and implications for foetal alcohol syndrome prevention in the Bergriver municipality in the Western Cape, South Africa. S Afr Geogr J 2014 Apr 14;96(2):153-165. [doi: 10.1080/03736245.2014.901186]
44. Bader MD, Ailshire JA, Morenoff JD, House JS. Measurement of the local food environment: a comparison of existing data sources. Am J Epidemiol 2010 Mar 1;171(5):609-617 [FREE Full text] [doi: 10.1093/aje/kwp419] [Medline: 20123688]
45. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med 2016 Jun;15(2):155-163 [FREE Full text] [doi: 10.1016/j.jcm.2016.02.012] [Medline: 27330520]
46. Haardoerfer R. Power and bias in hierarchical linear growth models: more measurements for fewer people (forthcoming). 2010. URL: https://pdfs.semanticscholar.org/bbdc/bd009f4ebda2399007bbd37e6941322f1318.pdf
47. Phillips CB, Engelberg JK, Geremia CM, Zhu W, Kurka JM, Cain KL, et al. Online versus in-person comparison of microscale audit of pedestrian streetscapes (MAPS) assessments: reliability of alternate methods. Int J Health Geogr 2017 Aug 4;16(1):27 [FREE Full text] [doi: 10.1186/s12942-017-0101-0] [Medline: 28778205]
48. Harding JP, Rosenblatt E, Yao VW. The contagion effect of foreclosed properties. J Urban Econ 2009 Nov;66(3):164-178. [doi: 10.1016/j.jue.2009.07.003]

Abbreviations

ICC: intraclass correlation
IQR: interquartile range

Edited by G Eysenbach; submitted 03.06.19; peer-reviewed by K Tamura, J Mennis, R Mathews; comments to author 28.06.19; revised version received 31.07.19; accepted 31.07.19; published 04.10.19.

Please cite as:
Crawford ND, Haardöerfer R, Cooper H, McKinnon I, Jones-Harrell C, Ballard A, von Hellens SS, Young A
Characterizing the Rural Opioid Use Environment in Kentucky Using Google Earth: Virtual Audit
J Med Internet Res 2019;21(10):e14923
URL: https://www.jmir.org/2019/10/e14923
doi:10.2196/14923
PMID:31588903

Original Paper

Tailored Web-Based Information for Younger and Older Patients with Cancer: Randomized Controlled Trial of a Preparatory Educational Intervention on Patient Outcomes

Minh Hao Nguyen[1,2], PhD; Ellen M A Smets[3], PhD; Nadine Bol[1,4], PhD; Eugène F Loos[1], PhD; Hanneke W M van Laarhoven[5], MD, PhD; Debby Geijsen[6], MD; Mark I van Berge Henegouwen[7], MD, PhD; Kristien M A J Tytgat[8], MD, PhD; Julia C M van Weert[1], PhD

[1]Amsterdam School of Communication Research, University of Amsterdam, Amsterdam, Netherlands

[2]Department of Communication and Media Research (IKMZ), University of Zurich, Zurich, Switzerland

[3]Department of Medical Psychology, Amsterdam University Medical Centers, University of Amsterdam, Amsterdam, Netherlands

[4]Department of Communication and Cognition, Tilburg University, Tilburg, Netherlands

[5]Department of Medical Oncology, Cancer Center Amsterdam, Amsterdam University Medical Centers, University of Amsterdam, Amsterdam, Netherlands

[6]Department of Radiation Oncology, Amsterdam University Medical Centers, University of Amsterdam, Amsterdam, Netherlands

[7]Department of Surgery, Cancer Center Amsterdam, Amsterdam University Medical Centers, University of Amsterdam, Amsterdam, Netherlands

[8]Department of Gastroenterology, Amsterdam University Medical Centers, University of Amsterdam, Amsterdam, Netherlands

Corresponding Author:
Minh Hao Nguyen, PhD
Amsterdam School of Communication Research
University of Amsterdam
Nieuwe Achtergracht 166, 1018 WV
Amsterdam
Netherlands
Phone: 31 20 525 368
Email: mh.nguyen@ikmz.uzh.ch

Abstract

Background: Many patients with cancer, including older patients (aged ≥65 years), consult the Web to prepare for their doctor's visit. In particular, older patients have varying needs regarding the mode in which information is presented (eg, via textual, visual, or audiovisual modes) owing to age-related sensory (eg, impaired vision and hearing) and cognitive decline (eg, reduced processing speed). Therefore, Web-based information targeted at older patient populations is likely to be used and processed more effectively, and evaluated more positively, when tailored to age-related capabilities and preferences. This, in turn, may benefit patient outcomes.

Objective: This randomized controlled trial tested the effects of a Web-based tailored educational intervention among newly diagnosed younger (<65 years) and older (≥65 years) patients with cancer. We compared the intervention group who viewed a mode-tailored website (ie, enabling patients to tailor information using textual, visual, and audiovisual modes) with 3 control groups view a nontailored website (ie, text only, text with images, and text with videos). We examined website experience outcomes (ie, website satisfaction, website involvement, knowledge, anxiety, and communication self-efficacy) and consultation experience outcomes (ie, question asking during consultation, anxiety, and information recall).

Methods: Patients from a multidisciplinary outpatient clinic (N=232) viewed a mode-tailored or nontailored website as preparation before their hospital consultations to discuss diagnosis and treatment. Data were collected before (T1), during (T2), and after (T3) visitation. Website experience outcomes were assessed with questionnaires (T1). Patients' question asking was coded from videotaped consultations, and anxiety was assessed through a questionnaire (T2). Telephone interviews were conducted to assess knowledge acquired from the website before (T1) and after consultation (T3), and information recall from the consultation (T3).

Results: The preparatory website was well used across all conditions (mean 34 min). Younger patients viewing the mode-tailored website were more satisfied before consultation (P=.02) and reported lower anxiety after consultation (P=.046; vs text only). This pattern was not found in older patients. Mode tailoring yielded no other significant differences in patient outcomes. Regression

analyses showed that website involvement (beta=.15; *P*=.03) and, to a lesser extent, website satisfaction (beta=.15; *P*=.05) positively associated with knowledge before consultation (T1). In turn, higher knowledge before consultation (beta=.39; *P*<.001), together with time on the website (beta=.21; *P*=.002; T1), predicted information recall from consultations (T3). Patients with higher knowledge before consultation (T1) also reported higher knowledge from the website afterward (T3; beta=.22; *P*=.003).

Conclusions: Offering preparatory online information before consultations benefits information processing and patient outcomes of both younger and older newly diagnosed patients with cancer. Younger patients benefit even more when information is offered in a mode-tailored manner. We discuss the theoretical, methodological, and practical implications for patient-provider communication research in an electronic health era.

Clinical Trial: Netherlands Trial Register NTR5904; https://www.trialregister.nl/trial/5750

(J Med Internet Res 2019;21(10):e14407) doi:10.2196/14407

KEYWORDS

Web-based tailoring; internet; audiovisual media; patient education; cancer; aging; memory; anxiety; patient reported outcomes; patient participation; consultation; health communication; randomized controlled trial

Introduction

Background

Cancer often occurs in people of older age (≥65 years), and this number is expected to grow globally [1]. Older patients with cancer constitute the majority of the cancer patient population and are also most at risk for poor communication with health care providers owing to age-related declines, such as in cognitive (eg, working memory) and physical functioning (eg, vision loss, hearing loss, and comorbidity) [2]. In general, older patients are less likely to express their information needs or preferences and participate less actively during consultations [2,3]. Moreover, they generally experience lower self-efficacy in obtaining relevant information from their provider [4] and have more difficulty remembering information from consultations than younger patients [5,6]. Therefore, particularly, older patients could benefit from support by communicating with providers. This study aimed to investigate whether tailored online health information can provide such support to older and younger patients by examining the effects on patient-reported and observed outcomes, including website satisfaction, communication self-efficacy, anxiety, question asking during consultation, and information recall.

The information society of today is characterized by the availability of and relatively easy access to cancer information on the internet. For many older adults, besides their health care provider, the internet is one of the first preferred health information sources [7]. Online health information (eg, a hospital website) is often used to prepare for a doctor's visit [7] and may lead to better informed, more confident, and less anxious patients [8,9]. Moreover, the use of preparation tools can support patients to actively participate in consultations (eg, by asking questions) and process and recall information from their health care provider [10,11].

Unfortunately, many older patients experience difficulties in using online health information [12]. Although this problem could resolve itself as generations pass by and the digital divide closes, age-related sensory (eg, impaired vision and/or hearing) and cognitive decline (eg, reduced processing speed) remain a prominent reason preventing older adults from using the internet effectively [12,13]. Such age-related declines also explain why older adults have varying needs regarding *how* information should be presented, making it more challenging to develop user-friendly websites for this group [14,15]. Online health information distinguishes itself from traditional formats of health information (eg, print materials) because of its possibility to integrate different modalities (ie, modes), such as textual, visual, and/or audiovisual information. What is particularly relevant for older patients is that these information modes can be tailored to match individual preferences and abilities (eg, age-related factors) and thus facilitate information processing [16]. *Mode tailoring* refers to the possibility of individuals to adapt the modality of online information presentation to their preferences, using textual, visual, and audiovisual information [17]. Recent experimental research showed positive effects on the evaluation, processing, and recall of cancer-related information when people were able to self-tailor the mode of presentation on a health website, especially among older adults [17,18]. Hence, mode tailoring is a particularly promising strategy to optimize online health information for the older population.

This study extends this experimental mode tailoring research to a clinical population of newly diagnosed patients with cancer who viewed a previsit website to prepare for their hospital consultations to discuss diagnosis and treatment planning. In a randomized controlled trial (RCT), we investigated both pre- and postvisit effects of exposure to a previsit website that can be tailored to patients' information mode preferences (by self-selecting text, images, and/or videos) compared with exposure to standardized, nontailored websites (with either text only, text with images, or text with videos). First, we examined the effects of mode tailoring on website experience outcomes before the consultation (T1), including patients' website involvement, satisfaction with the website, anxiety, self-efficacy in communicating with the provider, and knowledge. We also investigated whether mode tailoring effects extend to the consultation and beyond. Consultation experience outcomes include patients' question-asking behavior during consultation and anxiety (T2). Additionally, we considered knowledge gained from the website and information recall after the consultation (T3). Second, for all outcomes, we investigated how these effects differ between younger and older patients. Third, across all patients, we investigated how website experiences predict knowledge before the consultation (T1) and how website

experiences and consultation experiences predict knowledge and information recall after the consultation (T3). By looking into the interplay between online health information provision and offline patient-provider communication in the cancer context, this study has provided insights for both practice and theory regarding patient-provider communication in an electronic health (eHealth) era.

Mode Tailoring: Catering to Older Patients' Motivation and Ability

The elaboration likelihood model (ELM) and the limited capacity model of motivated mediated message processing (LC4MP) state that information processing is highly dependent on an individual's motivation (eg, attention) and ability (eg, cognitive resources) to process information [19,20]. Older adults often see themselves as less able and are less motivated to use online health information [21]. Moreover, many older adults who go online for health information are left unsatisfied [22]. A partial explanation is that many available health websites insufficiently consider age-related factors in their design [12,23]. Providing different information modes (eg, via text, visuals, and videos) in a tailored manner can increase both the *motivation* and *ability* to use and process online health information and may, therefore, be especially relevant for older users. For instance, when the mode of presentation matches with an individual's *preference* for how to consume online health information, this is likely to increase their *motivation* to attend to the information. Additionally, tailoring the mode of information presentation caters to differences in individual processing styles and abilities—including age-related declines in vision, hearing, and cognition—which *enables* individuals to process the information better. Thus, when online health information is tailored to individual mode preferences, these preconditions (ie, motivation and ability) for successful processing are considered more optimal. Consequently, mode-tailored information has a greater likelihood to reach and affect patients than nontailored information, especially older patients. In the following sections, we have discussed the expected benefits of mode-tailored online health information for patient-reported outcomes surrounding a hospital visit in younger and older patients.

Effects on Website Experience Outcomes Before Consultation: Involvement, Satisfaction, Anxiety, Communication Self-Efficacy, and Knowledge

A (potential) diagnosis of cancer typically involves high levels of anxiety [24], which can hinder patients' ability to process and remember information provided by their provider [18]. Although it can be overwhelming to receive information related to the disease [25], patients have a high need for information during this uncertain phase [26]. Providing patients with mode-tailored information might enable them to absorb the information in a dosed manner (eg, by reading the text first and saving a video for later) [27]. As mode-tailored information is more accessible to patients, it is expected to be evaluated and processed better too [17,18]. Furthermore, viewing tailored online information before a hospital visit is likely to decrease patients' anxiety, as they are better informed and prepared for what can be expected [28]. Similarly, the use of tailored preparatory information might increase self-efficacy in communicating with the provider [29,30]. Given that mode tailoring is anticipated to cater to age-related declines, we expect that older patients will benefit more from a mode-tailored website than younger patients.

- H1: Exposure to a mode-tailored preparatory website (vs nontailored websites) will affect patients' website experience outcomes before a consultation (T1), including enhanced website involvement (H1a); enhanced website satisfaction (H1b); decreased anxiety (H1c); enhanced communication self-efficacy (H1d); and improved knowledge (H1e).
- H2: These effects will be stronger for older patients (≥65 years) than for younger patients (<65 years) with regard to website involvement (H2a), website satisfaction (H2b), anxiety (H2c), communication self-efficacy (H2d), and knowledge (H2e).

Effects on Consultation Experience Outcomes: Question Asking and Anxiety

The abovementioned effects of mode-tailored online information might also extend to the consultation and beyond (eg, patients' question asking and anxiety). Combining tailored preparatory information and interpersonal patient-provider communication can reinforce each other's effectiveness [31]. For instance, viewing preparatory information may make patients aware of topics of information they would like to know more about or validate with their health care provider, causing them to be more actively involved during consultations by asking questions [30,32]. Alternatively, patients viewing preparatory information before consultation may feel better informed and prepared for their visit, resulting in fewer questions asked during consultation [29]. Regardless, providing preparatory information in a tailored manner could strengthen effects in both directions (ie, more or less questions). As it is unclear how viewing mode-tailored online information before consultation would affect patients' question asking and how this differs by age, the following research questions were formulated:

> *RQ1: Does exposure to a mode-tailored preparatory website (vs nontailored websites) lead to more or less questions asked by patients during consultation (T2)?*
>
> *RQ2: To what extent does the relation between exposure to a mode-tailored preparatory website and patients' question asking during consultation differ between younger (<65 years) and older patients (≥65 years)?*

Preparatory information can play a key role in limiting anxiety during cancer consultations, perhaps even more so for patients who tend to avoid information [24]. Bronner et al found that patients with cancer characterized by a monitoring coping style, that is, *information seekers*, became less anxious from pre- to postconsultation after receiving their diagnosis and treatment plan [24]. The opposite relation was found for patients identified more as *information avoiders*; this group became more anxious from pre- to postconsultation, especially when receiving bad news [24]. A possible explanation is that information seekers had already searched for information before their consultation and were prepared for the worst scenario. Thus, when hearing

their diagnosis and treatment advice, they felt relieved when hearing relatively *good* news or they were more prepared for bad news. This is in contrast with the *less prepared* information avoiders who became more distressed after their consultation when receiving bad news. Therefore, we expect that the use of preparatory information before hospital visits, especially when tailored, decreases anxiety immediately after consultation.

- H3: Exposure to a mode-tailored preparatory website (vs nontailored websites) will decrease anxiety immediately after consultation (T2).

Additionally, we expect the effect of mode-tailored preparatory information on anxiety to be especially visible in older patients. The socioemotional selectivity theory posits that as people age, goals associated with emotional meaning and well-being become more salient, whereas knowledge-related goals to prepare for future events become less important [33]. Consequently, older adults generally process information in such a way that it helps them regulate their emotions (eg, putting them at ease). Older patients may perceive more emotional gratification from information presented in visual and audiovisual modes because these modes often include more vivid and obvious personal elements (eg, a patient video) that appeal more to their emotion-oriented preferences and needs. For instance, previous research has shown that older adults often prefer visual and audiovisual information [15,34], and such information modes have been found to increase feelings of emotional support from online cancer-related information compared with text in older people [35,36]. However, studies have also shown high variability in older adults' information mode preferences [12,37]. As such, providing older patients with the option to select their preferred information modes, including visual and audiovisual elements, is more likely to fulfil their emotional and informational needs, thereby limiting their anxiety.

- H4: The effect of mode tailoring on anxiety immediately after consultation will be stronger for older patients (≥65 years) than for younger patients (<65 years).

Effects on Knowledge and Recall of Information After the Consultation

Consulting online information before consultations might also improve knowledge from online information and information recall after the consultation. For instance, when patients are already informed about several topics before consultations, this could prime patients' attention to these and related topics when being discussed by the provider during consultations (ie, a repetition effect) [38]. Additionally, being informed and knowing what to expect beforehand could leave patients with more cognitive capacity to attend to new information that they receive during consultations. In other words, providing information in a dosed manner over multiple occasions allows patients to process important information at a slower pace which may benefit information recall [39]. As older patients have more difficulty in remembering medical information, it is expected that they will benefit relatively more from mode-tailored preparatory information than younger patients.

- H5: Exposure to a mode-tailored preparatory website (vs nontailored websites) will improve knowledge from the website and information recall from the consultation (T3).
- H6: The effect on knowledge from the website and information recall from the consultation will be stronger for older patients (≥65 years) than for younger patients (<65 years).

Motivation- and Ability-Related Factors Explaining Patients' Information Processing

Besides the main effects of mode tailoring on patient-reported outcomes before, during, and after the consultations, different *website experience outcomes* (eg, website involvement) may independently explain knowledge before the consultation (T1) and, together with *consultation experience outcomes* (T2; eg, question asking and anxiety), predict knowledge from the website and information recall from the consultation (T3). The different processes explaining knowledge acquisition and information recall can be related to a patient's *motivation* (eg, website involvement) or *ability* (eg, communication self-efficacy) to process information. Although ELM and LC4MP are useful frameworks in understanding how mode tailoring can enhance *motivation* and *ability* to process information, how these processes translate to specific variables explaining knowledge and recall of information in the cancer context has only been briefly explored [5,6,40]. For instance, Bol et al used ELM and LC4MP to identify motivation- and ability-related factors in the literature deemed relevant for processing of online cancer information [40]. They identified website involvement and website satisfaction as *website experience outcomes* positively associated with information recall, whereas perceived cognitive load was negatively related to information recall. However, Bol et al did not examine which *consultation experience outcomes* contribute to effective information processing in patient-provider encounters [40]. Thus, in addition to addressing the value of mode tailoring, this study sought to gain insight into which *website experience outcomes* and *consultation experience outcomes* explain the benefits of preparatory online information on knowledge acquisition from websites and information recall from consultations in patients with cancer, as well as how these concepts relate to each other over time. By doing so, we inform future research relying on theories such as the ELM and LC4MP to understand which specific *motivation*- and *ability*-related factors play a role in how online and offline cancer information is being processed. We explored the following research questions:

> *RQ3: Which website experience outcomes (eg, website use, website involvement, and anxiety) predict knowledge from the website before a consultation (T1)?*
>
> *RQ4: Which website experience outcomes (T1) and which consultation experience outcomes (ie, question asking and anxiety; T2) predict knowledge from the website and information recall after the consultation (T3)?*

Methods

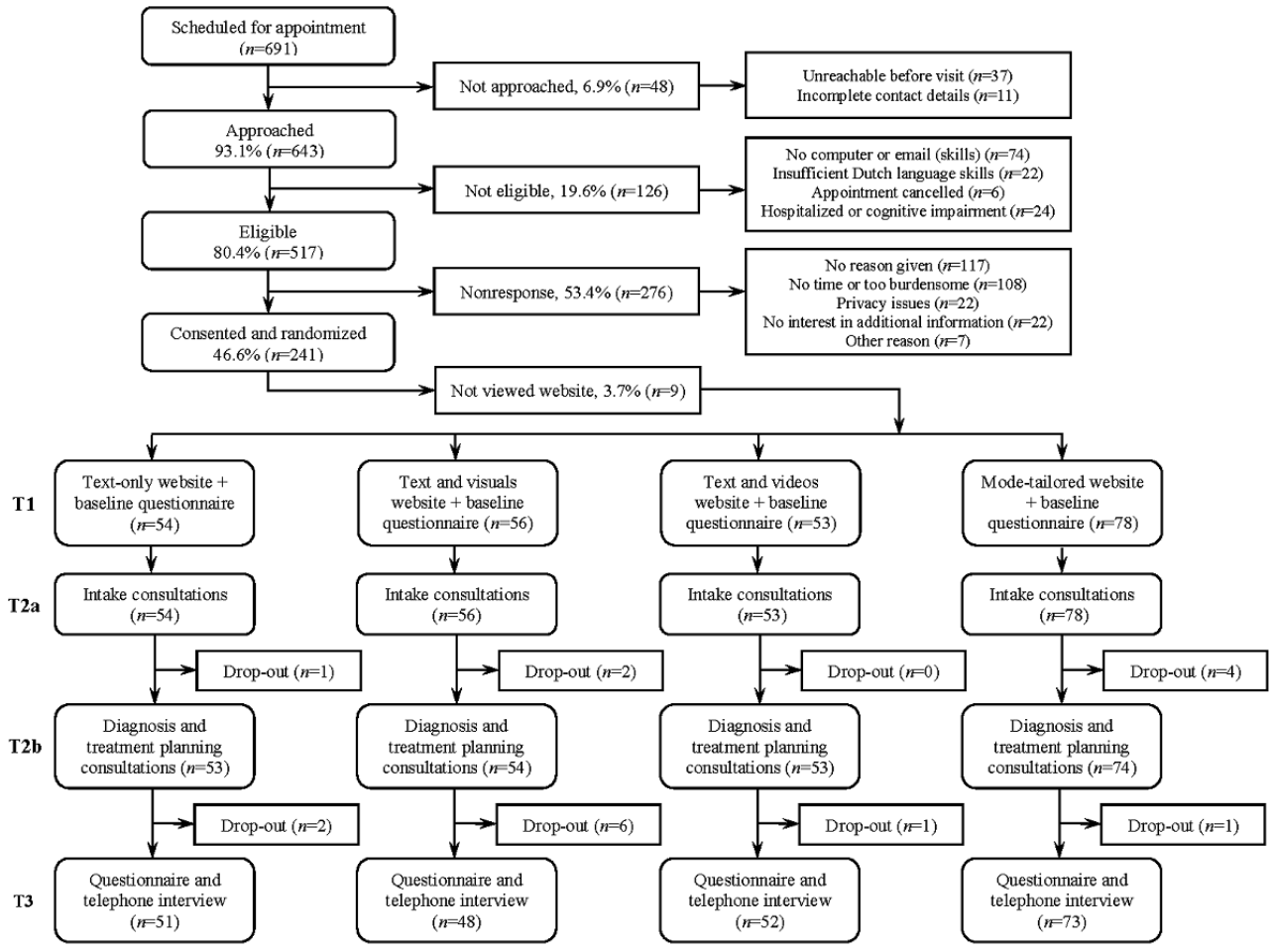
Design

An RCT was conducted to compare the effectiveness of a mode-tailored website (with options to choose text, visuals, and/or videos) with 3 standardized, nontailored versions with text only, text with visuals, and text with videos. Patients were stratified into a younger group (<65 years) and older group (≥65 years) and randomly assigned to view one of the 4 website versions. An age cut-off of 65 years was selected, as cancer more frequently occurs in adults above this age [1] and as similar studies on older patients and online health information also used this cut-off age [12,35,36]. An *a priori* power analysis based on an analysis of variance (ANOVA) with 8 groups (condition × age group) revealed that a sample size of 237 was needed to detect a medium-sized effect (Cohen *f*=0.25) with an observed power of 0.80 and an alpha level of .05. The study was approved by the medical ethical review board of the Amsterdam University Medical Center (reference number: W13_053 #13.17.0069) and the ethics committee of the Amsterdam School of Communication Research (reference number: 2014-CW-110).

Participants

Participants were patients who were suspected of having colorectal, stomach, or esophageal malignancies or had received a preliminary cancer diagnosis (but were awaiting information on the tumor stage based on additional imaging or came for a second opinion) who were referred to the Gastro-Intestinal Oncological Centre Amsterdam (GIOCA). Patients were recruited from December 2015 through September 2018. The GIOCA is an academic multidisciplinary outpatient clinic in the Netherlands that specializes in fast-track diagnosis and treatment planning within a day [41], referred to as the *GIOCA day*. During the study, 691 patients visited the GIOCA; of these, 643 were successfully approached by telephone 1 to 5 days before their visit, depending on the day they received their referral. We informed patients about the purpose of the study (ie, to gain insight into information provision to patients with cancer) and offered them access to a website containing relevant information about GIOCA's procedures, which could help them prepare for their visit. Of the 517 patients who had internet access and wanted to receive an email with access to the website, 241 consented to participate. As 9 of the included patients did not use the website, a total of 232 patients were included in the final analyses. Patients most often declined participation because they had no time or found it too burdensome. Only 8.0% (22/276) of declining patients explicitly mentioned that recording their consultation was a breach of their privacy, and only 8.0% (22/276) had no interested in additional (online) information. A nonresponse analysis revealed no differences between participating and nonparticipating patients in age, t_{689}=1.52; *P*=.13, and gender, χ^2_1=3.2; *P*=.07. An overview of participant inclusion, reasons for nonresponse, randomization procedure, and dropout rates is presented in Figure 1.

Figure 1. Flowchart of participant recruitment and drop-out.



Website Intervention

A total of 4 website versions were developed containing the same information but presented in different modalities (via text, images, and/or patient videos). The nontailored website versions contained either text only, text with images, or text with videos. The text with videos version contained 6 videos featuring patients who narrated the textual information on the website (ranging from 1:30 to 3:00 min in length). As the images and videos were based on the textual content, they offered similar information. The information on the nontailored websites was offered in a standardized manner and could not be adapted by patients. The mode-tailored website version allowed patients to self-tailor the information presentation to their preferred mode at any moment during viewing. We did not include a nontailored version with all modalities, as we previously found that too much information on one Webpage can be detrimental for patient outcomes [18].

The website contained different pages with information about the fast-track clinic, how to prepare for consultations, and when to contact the clinic. Furthermore, the website contained information about the conditions (colorectal, stomach, or esophageal cancer), medical tests, treatment options, and practical information, such as a list of health care providers, frequently asked questions, and contact and location information. The content on the websites was similar for both patients with colorectal cancer and patients with stomach and esophageal cancer, except for the information concerning the condition and treatment options. Details on the development and content of these websites are published elsewhere [27]. Examples of the mode-tailored website and the nontailored websites are given in Figures 2 and 3.

Figure 2. Example of the mode-tailored website with all modes switched on (text, images, videos).

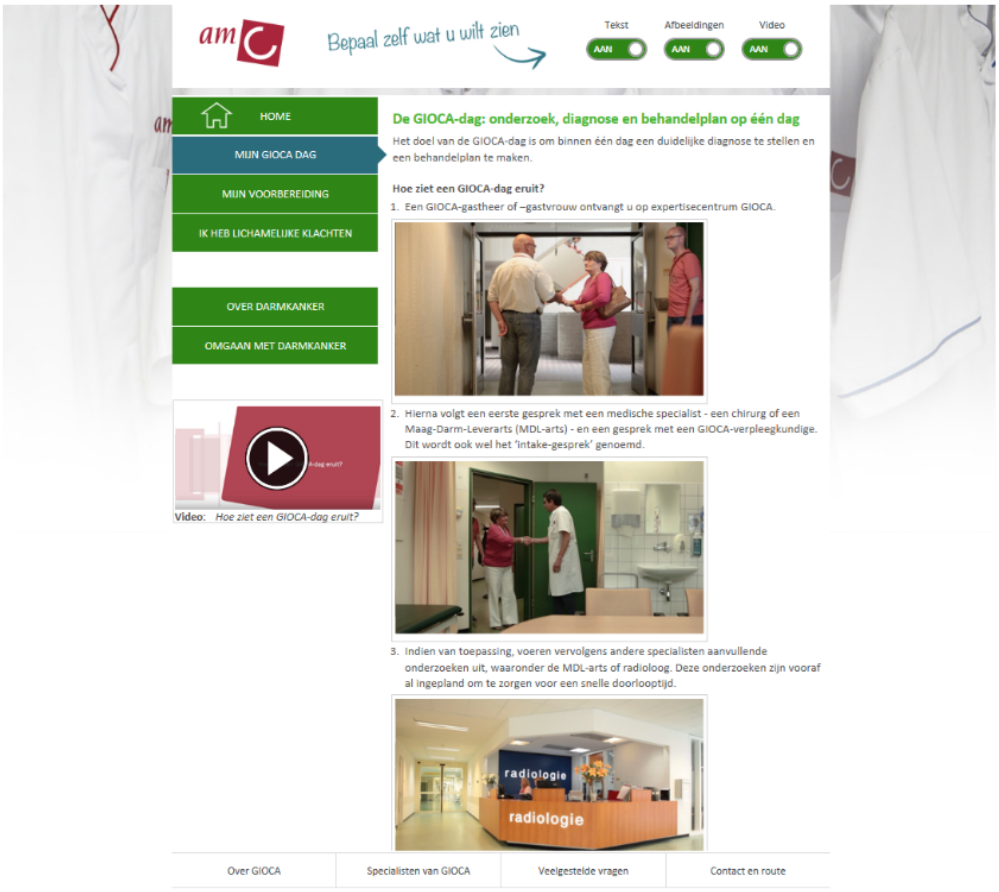
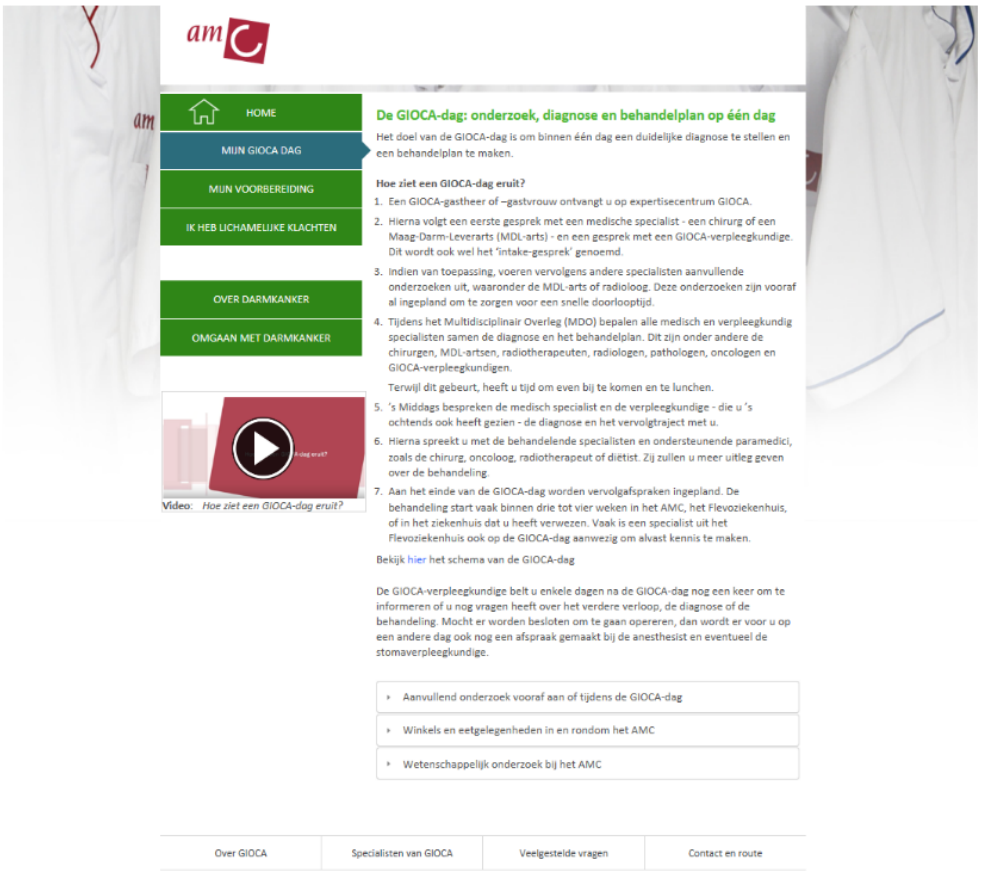


Figure 3. Example of a non-tailored website with text and video.



Procedure

Consenting patients were stratified by age group and then randomized to one of the 4 conditions using randomization software. After the first telephone contact (1-5 days before visit), patients received a link to one of the 4 website versions by email from the study coordinator. Patients were not aware that there were other website versions than the one they received. Patients were free to use the website as they wished (how often, how long, and which pages). After viewing the website, consenting patients completed an online questionnaire to record website experience outcomes and background variables such as sociodemographic information, health background information, and information preference characteristics (T1a). One day before their visit to the clinic, patients' knowledge acquired from the website was assessed by telephone (T1b). On the day of the patients' visit to the clinic, a research assistant was present in the hospital to video record all consultations to assess question asking during consultations (T2). The fast-track program (GIOCA day) started with 2 intake consultations (medical specialist and nurse) to evaluate symptoms and medical history (T2). At noon, a multidisciplinary team discussed the diagnosis and formulated a treatment plan. In the afternoon, the diagnosis and treatment advice were discussed with the patient by the physician and nurse who conducted the intake consultations (T2). Depending on the treatment plan, patients also visited a surgeon, oncologist, or radiation oncologist on the same day to discuss treatment details (T2). Patients usually had 4 to 7 consultations during the GIOCA day, all of which were video recorded for this study. Immediately after the last consultation, a paper questionnaire was used to measure anxiety (T2). Patients were contacted by telephone within 36 to 48 hours after their visit to assess their knowledge from the website and information recall from the consultation (T3).

We took extensive measures to limit the overall participant burden. We had a trained research assistant who informed patients at the start of the study by telephone to ensure they knew exactly what the participation would entail, including how their privacy was safeguarded. Then, we sent out an information letter over email and contacted patients after 2 days again. Every patient was in contact with 1 research assistant who conducted telephone interviews before and after consultation and accompanied the patient in the hospital for data collection (handing out the questionnaires at the right time and installing the video recorder). This ensured that both patients and providers dealt with 1 person only and did not have to memorize when to fill out which questionnaire or turn on the video recording.

Perceived Level of Tailoring of the Website

We measured to what extent patients felt the website was tailored to their situation with 2 items. The items comprised "the way I viewed information on the website (via text; text with images; text with video; text, images, and/or video) corresponded to my preference to receive health information" and "the presentation of the information on the GIOCA website was tailor-made for me." The answer options ranged from 1 (totally disagree) to 7 (totally agree), with higher scores indicating higher levels of perceived tailoring (Pearson r=0.74; P<.001).

Website Experience Outcomes (T1)

Website use patterns were recorded using a built-in Web tracker that logged every action on the website, including the number of clicks, time spent on each page, video viewing behavior, mode selections, and number of visits.

Website involvement was measured with 5 items, including "I was highly involved in evaluating the website" and "I carefully viewed the information on the website" [42]. Answer options ranged from 1 (totally disagree) to 7 (totally agree), of which mean scores were calculated (alpha=.81).

The 10-item version of the *Website Satisfaction Scale* [18,43] was used to assess the degree to which patients were satisfied with the (1) attractiveness of the website (3 items, eg, *the website looks nice*; alpha=.86), (2) comprehensibility of the website (3 items, eg, "the website is understandable"; alpha=.97), and emotional support from the website (4 items, eg "the website helps me with my emotions"; alpha=.92). Answer options ranged from 1 (totally disagree) to 7 (totally agree), of which mean scores were calculated.

Communication self-efficacy was measured with the short-form Perceived Efficacy in Patient-Physician Interactions questionnaire [44]. A total of 5 questions assessed the patient's confidence in their ability to communicate with their provider on a scale from 1 (very confident) to 5 (not confident at all). A sum score was calculated (range 5-25; alpha=.88).

Patients' current state of *anxiety* was measured with the 6-item version of the State-Trait Anxiety Inventory (STAI-6) [45]. Patients rated whether they experienced the presence (tense, upset, or worried) or absence (calm, relaxed, or content) of anxiety from 1 (not at all) to 4 (very much so). On the basis of the guidelines, the scores were recoded to range from 20 to 80, with scores >44 indicating high anxiety (alpha=.80) [46].

Knowledge acquired from the website was measured using the protocol of the Netherlands Patient Information Recall Questionnaire (NPIRQ) via telephone interviews [6,47]. We asked patients 12 standardized open questions based on the website content (eg, about the goal and course of the fast-track program and which medical specialists they will see). On the basis of a predeveloped code book, for 9 questions, patients could score 2 points, and for 3 questions, the correct answers contained fewer or more elements, and thus, these questions accounted for 1, 2.5, and 3 points. Thus, the maximum knowledge score at T1 for 12 questions (9×2 + 1×1 + 1×2.5 + 1×3) was 24.5. A standardized score was calculated by taking the percentage correctly answered according to the NPIRQ guidelines. The first author (MHN) coded all the answers. Additionally, answers from 14 patients were double coded by a second coder (FY), showing good intercoder reliability (mean kappa=0.737; *P*=<.001).

Consultation Experience Outcomes (T2)

Anxiety immediately after consultation was measured in the same way as before the consultation (STAI-6; mean 41.45, *SD* 11.76; Cronbach alpha=.81).

To code the *question-asking behavior*, a codebook was developed based on earlier work by Zandbelt et al [48]. All questions during consultation related to (1) medical, (2) practical, and (3) paramedical information were coded. Questions on medical topics were about the patients' disease, treatment (options), complications, and side effects. Questions on practical topics were about logistics of treatment and follow-up appointments. Questions on paramedical information were about psychosocial topics and consequences for daily life. Questions that were unrelated to the patient's condition (eg, the weather and holiday) were not coded. All questions were summed into one total score. A total of 16 consultations were double coded, revealing good intercoder reliability (Krippendorff alpha=.951).

Recall of Information and Knowledge After the Consultation (T3)

Information recall from the consultation and *knowledge* from the website were measured with the NPIRQ (similar to measurement of knowledge before the consultation; T1). Regarding *information recall* from the consultation, each participant was asked 13 standardized open questions (eg, about the proposed treatment plan, logistic planning of treatment, possible risks and side effects of treatment, and recommendations for daily life). To improve the validity of the recall measure, a maximum of 5 additional open questions were formulated, tailored to each patient's videotaped consultations (eg, about details of treatment and additional medical tests). The correct answers were derived from the videotaped consultations. Each answer as provided by the participant during the interview was scored as not recalled (0), partially recalled (1), and completely recalled (2) based on a predeveloped code book. In theory, patients could receive a maximum of 18 questions (13 standardized + 5 tailored questions). However, as the content of consultations varied between patients, not all standardized open questions were applicable to all patients. Similar to T1, a standardized score was calculated by taking the percentage of correctly recalled information, based on the patient's total sum score (1-23) and the maximum obtainable recall score (range 4-34; mean 56.37, *SD* 15.21).

The website contained information on 10 of the 13 standardized open questions asked. Hence, a separate *knowledge* score about topics on the website was calculated from these 10 questions. For this knowledge score, the website content was used as a guideline to score patients' answers (similar to knowledge at T1). Again, a standardized score was calculated by taking the percentage correctly answered, based on the total sum score (0-9) and maximum obtainable score (range 10-19; *mean* 15.05, *SD* 11.76). The first author (MHN) coded all answers. A second coder (MA) coded answers from 14 participants from a different dataset with the exact same code book (mean kappa=0.816; *P*<.001) [6].

Background Variables

Sociodemographic Information

Sociodemographic information included age, gender, and education level. Education level was divided into lower (ie, primary education, general secondary education, and middle vocational education) and higher education level (ie, higher vocational education and university).

Health Background Information

Health background information included cancer type (coded as colorectal cancer=0 and esophageal/stomach cancer=1) and whether patients came in for a second opinion (no=0 and yes=1). The treatment goal (palliative or unclear=0 and curative=1) was derived from the medical file. Health literacy was measured with the comprehension test of the Short Assessment of Health Literacy in Dutch that consists of 22 health-related words (eg, biopsy, ventricle, and palliative) [49]. For each word, people were asked to select the correct meaning out of 3 multiple choice options or an "I don't know" option. The sum score of correct answers reflects their health literacy level and could range between 0 and 22. Patients' frailty (ie, functioning in the physical, cognitive, social, and psychosocial domain) was assessed with the 15-item Groningen Frailty Indicator [50]. The quality of life was measured with 2 items from the European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire [51] with answer options ranging from 1 (very bad) to 7 (excellent; Pearson r=0.75; P<.001).

Information-Seeking Characteristics

Internet use was measured in hours per week. Monitoring coping style refers to the degree to which patients seek information in a threatening medical situation. This was assessed with an adapted version of the Threatening Medical Situation Inventory (eg, "I intend to get as much information as possible about my treatment") [24,52], using 3 items with answer options from 1 (not applicable to me at all) to 5 (very applicable to me; Cronbach alpha=.83). We assessed information preference with an adapted item from the Information Satisfaction Questionnaire [53], asking whether patients prefer to receive (1) "as much information as possible, both positive and negative," (2) "as much information as possible, both positive and negative, but bit by bit," (3) "not much information," and (4) "only positive information". In conformity with previous research [54], the items were dichotomized by merging category (2), (3), and (4) into "not all information (at once)" (0) versus "as much information as possible, both positive and negative" (1). Finally, we assessed whether patients received information about the clinic from other sources (eg, health care providers and brochures) besides our website (no=0 and yes=1).

Statistical Analyses

Chi-square tests, *t* tests, and ANOVAs were conducted to check for unequal distribution of background variables over conditions. Descriptive analyses were used to explore patterns of website use. Main and interaction effects of mode tailoring (H1-H6, RQ1-2) were tested with ANOVAs. Additional simple effects analyses were conducted to examine differences between conditions within age groups. The significance level was set at P<.05. To test which website experience outcomes predicted knowledge before the consultation (T1) and how these, together with consultation experience outcomes (T2), predicted information recall from the consultation and knowledge from the website after the consultation (T3; RQ3-4), 3 multistage linear regression models were estimated. All analyses started with a baseline model (Model 1) of individual background variables (with age as a continuous variable). Website experience outcomes were added as predictors in Model 2. For information recall from the consultation and knowledge from the website at T3, relevant consultation experience outcomes (T2) were included as predictors in Model 3. To reduce the number of predictors, only variables that were at least marginally correlated with knowledge or information recall were included in the models (P<.10). Assumptions of linearity, normality, homoscedasticity, independent errors (Durbin-Watson values for Model 1, 2, and 3 are 1.95, 1.89, and 1.75, respectively) and multicollinearity (variance inflation factor<10) were met for all variables. Standardized coefficients (betas) are reported for comparisons of predictive power.

Results

Patient Characteristics

Participating patients were aged, on average, 63.50 years (*SD* 9.06; range 36-86), with 46.1% (107/232) aged ≥65 years. The majority were male (68.1%, 158/232) and lived together with their spouse, children, or other family members (82.8%, 192/232). The majority were advised a curative treatment plan (73.2%, 170/232), whereas 13.4% (31/232) entered a palliative trajectory, and 13.4% (31/232) were scheduled for additional imaging studies to formulate a clear diagnosis and treatment plan. Of the 232 participating patients, 74 viewed the mode-tailored website, 53 viewed the text-only website, 54 viewed the text with images website, and 51 viewed the text with video website. Background information of the patients is given in Table 1.

Table 1. Patient background characteristics.

Background variables[a]	Older patients (n=107)	Younger patients (n=125)	All patients (n=232)	Total (N)[b]
Sociodemographic information				
Age (years), mean (SD)	71.44 (4.23)	56.81 (6.18)	63.50 (9.06)	232
Gender				
Male, n (%)	77 (72.6)	81 (64.3)	158 (68.1)	232
Female, n (%)	29 (27.4)	45 (35.7)	74 (31.9)	232
Education level				
Lower, n (%)	71 (67.0)	75 (59.5)	146 (62.9)	231
Higher, n (%)	35 (33.0)	50 (39.7)	85 (37.1)	231
Health background information				
Cancer type [c]				
Colorectal, n (%)	77 (72.6)	110 (87.3)	187 (77.9)	232
Esophageal/stomach, n (%)	29 (27.4)	16 (12.7)	45 (22.1)	232
Second opinion				
No, n (%)	90 (84.9)	99 (78.6)	189 (81.5)	232
Yes, n (%)	16 (15.1)	27 (21.4)	43 (18.5)	232
Treatment goal				
Palliative, n (%)	13 (12.3)	18 (14.3)	31 (13.4)	232
Curative, n (%)	81 (76.4)	89 (70.6)	170 (73.2)	232
Unclear, n (%)	12 (11.3)	19 (15.1)	31 (13.4)	232
Health literacy[d], mean (SD)	16.50 (4.99)	16.82 (4.38)	16.66 (4.66)	182
Frailty[e], mean (SD)	2.19 (1.86)	2.69 (2.05)	2.46 (1.98)	182
Quality of life[f], mean (SD)	5.24 (1.17)	4.99 (1.36)	5.11 (1.28)	229
Information characteristics				
Internet use[b,g], mean (SD)	12.02 (10.30)	17.98 (17.74)	15.27 (15.10)	229
Information coping style[h], mean (SD)	3.74 (0.87)	3.76 (0.87)	3.75 (0.87)	229
Information preference				
Not all information, n (%)	25 (24.0)	31 (24.8)	56 (24.5)	229
As much information as possible, n (%)	79 (76.0)	94 (75.2)	173 (75.5)	229
Additional information received				
No, n (%)	30 (30.9)	38 (31.4)	68 (31.2)	218
Yes, n (%)	67 (69.1)	83 (68.6)	150 (68.8)	218

[a]No differences were found between conditions.

[b]*N* refers to the entire population under study and *n* refers to a sample population under study. Not all cells add up to 100% owing to missing data.

[c]Differs significantly between younger and older patients at $P<.01$.

[d]A higher score indicates higher levels of health literacy (maximum range: 0-22; reported range 0-22).

[e]A higher score indicates higher frailty (maximum range 1-15; reported range 0-10).

[f]A higher score indicates higher quality of life (maximum range 1-7; reported range 2-7).

[g]Measured in hours per week.

[h]A higher score indicates a higher information monitoring coping style (maximum range 1-5; reported range 1-5).

Perceived Level of Tailoring

Patients viewing the mode-tailored website version had equally high perceptions of the degree to which the information presentation was tailored to them as compared with those viewing the nontailored versions (*mean* 5.24, *SD* 1.22; $F_{3,225}=0.19$; $P=.91$; $\eta p^2=0.00$).

Website Use Patterns

Of the 232 participating patients, 74 viewed the mode-tailored website (31.9%), 53 viewed the text-only website (22.8%), 54 viewed the text with images website (23.3%), and 51 viewed the text with video website (23.0%). Patients spent an average of 34 min and 45 seconds on the website (*SD* 00:32:56; range 00:00:34-03:50:42). Patients who received the website a day before their visit did not spend less time on the website than those who received it earlier (t_{230}=1.79; *P*=.07) and this did not differ between conditions (χ^2_3=1.1; *P*=.76). The majority of patients (62.1%, 144/232) visited the website twice or more in the days before their visit (mean 2.78, *SD* 2.28; range 1-22). Patients mostly consulted information about the GIOCA day (90.9%, 211/232), how to prepare for their visit (86.6%, 201/232), their condition (colorectal, stomach, or esophageal cancer; 80.6%, 187/232), and how to deal with cancer in daily life (ie, nutrition, fatigue, and psychosocial care; 67.2%, 156/232), and with which symptoms to contact the hospital (65.1%, 151/232). Almost half of the patients viewed information on diagnostic tests (40.5%, 94/232) and frequently asked questions (48.3%, 112/232). Contact information (27.2%, 63/232), information about which medical specialists work at GIOCA (19.8%, 46/232), and information on additional websites (25.0%, 58/232) were least often consulted. An overview of website use patterns is presented in Table 2.

Videos were available for patients in the text with video condition and mode-tailored condition (total *n*=125). Of these patients, 41 (32.8%; $n_{tailored}$=18; n_{video}=23) watched a total of 96 videos on the website. Within the conditions, 28% (21/74) of patients in the tailored condition watched a video compared with 39% (20/51) in the text with video condition. Within age groups, 29% (16/54) of older patients watched a video compared with 35% (25/71) of younger patients. These differences were not significant. Most patients who watched a video, watched it almost completely (total video time: mean 00:06:52, *SD* 00:05:19, range 00:15-22:35). The majority of patients who watched videos, watched more than one (61%; 25/41).

Patients in the mode-tailored condition spent an average of 43:07 min on the website, compared with 30:59 min for patients in the text condition, 33:52 min for patients in the text with images condition, and 26:26 min for patients in the text with video condition. However, this difference was not significant ($F_{3,224}$=2.52; *P*=.06; ηp^2=0.03). There were no differences between age groups in terms of time spent on the website ($F_{3,224}$=0.00; *P*=.96; ηp^2=0.00). All patients in the mode-tailored condition chose at least text, but the majority supplemented this with additional images or videos spread over multiple visits to the website. The majority of patients (77%, 57/74) in the mode-tailored condition selected all 3 modalities (text, images, and video); 16% of patients (12/74) chose text and images; 1% of patients (4/74) chose text only; and only 1/74 patient chose text with video. Regarding the first time on the website, most patients first chose text (79%, 59/74); 13% of patients first chose images (10/74); and 6% of patients first chose video (5/74). During subsequent Web sessions, patients were more likely to choose images and video first. On average, patients took 01:15 min to select their first mode (*SD* 02:10; 39%, 29/74 <30 seconds, 66%, 50/74 <1 min, and 86.5%, 64/74 <2 min). Regarding the second visit, the first mode was selected on average at 29 seconds (*SD* 00:51; 84%, 32/38 <30 seconds, 94.7%, 36/38 <2 min).

Table 2. Patterns of website use.

Website use variables	Older patients (n=106)	Younger patients (n=126)	All patients (n=232)
Time spent on website (mm:ss), mean (SD)	34:27 (32:09)	35:00 (33:42)	34:45 (32:56)
Mode-tailored (31.9%, n=74)	41:35 (40:55)	44:14 (43:01)	43:07 (41:53)
Text-only (22.8%, n=53)	30:25 (39:34)	31:31 (22:11)	30:59 (25:49)
Text with images (23.3%, n=54)	34:09 (30:34)	33:35 (35:23)	33:52 (32:51)
Text with video (23.0%, n=51)	29:42 (21:55)	25:33 (20:15)	26:26 (20:55)
Web pages, n (%)			
The GIOCA[a]-day	90 (84.9)	121 (96)	211 (90.9)
Preparing for the GIOCA-day	84 (79.2)	117 (92.9)	201 (86.6)
Information about cancer types	84 (79.2)	103 (81.7)	187 (80.6)
Diagnostic tests	49 (46.2)	45 (35.7)	94 (40.5)
When to contact the hospital	62 (58.5)	89 (70.6)	151 (65.1)
Daily life recommendations	65 (61.3)	91 (72.2)	156 (67.2)
Additional relevant websites	23 (21.7)	35 (27.8)	58 (25.0)
Frequently asked questions	48 (45.3)	64 (50.8)	112 (48.3)
Medical specialists at GIOCA	23 (21.7)	23 (18.3)	46 (19.8)
Contact information	31 (29.2)	32 (25.4)	63 (27.2)
Watched at least one video[b], n (%)	16 (29.6)	25 (35.2)	41 (32.8)
Mode-tailored	9 (29.0)	12 (27.9)	21 (28.4)
Text with video	7 (30.4)	13 (46.4)	20 (39.2)
Number of videos watched[c], n (%)	2.50[d] (2.50)	2.24[e] (1.30)	2.34[d] (1.84)
Mode-tailored	3.00[d] (3.28)	2.50[e] (1.57)	2.71[d] (2.39)
Text with video	1.86[f] (0.69)	2.00[g] (1.00)	1.95[g] (0.89)
Number of mode actions[h], mean (SD)	2.97 (7.27)	2.72 (6.26)	2.84 (6.72)
Time until first mode (mm:ss)[h], mean (SD)	01:00 (00:48)	01:25 (02:45)	01:15 (02:10)
First mode ≤1 min (%)	64.5	67.4	66.2
First mode ≤2 min (%)	87.1	86	86.5
First mode ≤4 min (%)	100	95.3	97.3
First mode chosen[b,h], n (%)			
Text	27 (87.1)	32 (74.4)	59 (79.7)
Illustrations	4 (12.9)	6 (14.0)	10 (13.5)
Video	0 (0.0)	5 (11.6)	5 (6.8)
Mode combinations[b,h], n (%)			
All 3 modes	23 (74.2)	34 (79.1)	57 (77.0)
Text and illustrations	8 (25.8)	4 (9.3)	12 (16.2)
Text and video	0 (0.0)	1 (2.3)	1 (1.4)
Text only	0 (0.0)	4 (9.3)	4 (5.4)

[a]GIOCA: Gastro-Intestinal Oncological Centre Amsterdam

[b]Only applicable to patients viewing the mode-tailored (n=74) and text with video website (n=51).

[c]Only includes patients who watched at least one video.

[d]Range: 1-11.

[e]Range: 1-6.

[f]Range: 1-3.

[g]Range: 1-4.

[h]Only applicable to patients viewing the mode-tailored website (n=74).

Effects on Website Experience Outcomes Before Consultation (T1)

Table 3 shows the summary statistics of all outcomes. We hypothesized that exposure to a mode-tailored website (vs nontailored websites) would positively affect patients' website involvement, website satisfaction, anxiety, communication self-efficacy, and knowledge before the consultation (H1). Our data showed no significant differences between the conditions for these website experience outcomes. We also hypothesized differential effects of mode tailoring for younger and older patients, with stronger effects for older patients (H2). Although no significant interaction effects were present, a simple effects analysis revealed that, in contrast with our hypothesis, for younger patients the mode-tailored website (mean 5.12, *SD* 0.97; *P*=.02) and text with images website (mean 5.30, *SD* 0.91; *P*=.009) resulted in higher satisfaction with the attractiveness of the website compared with the text-only website (mean 4.46, *SD* 1.08). In general, older patients had lower knowledge levels (mean 22.70, *SD* 12.57) than younger patients (mean 30.16, *SD* 13.00; $F_{3,218}$=17.91; *P*<.001; ηp^2=0.08). Overall, the data showed no support for H1 and H2.

Table 3. Means and standard deviations of patient outcome variables.

Patient outcome variables	Mode-tailored, mean (SD)		Text-only, mean (SD)		Text with images, mean (SD)		Text with video, mean (SD)		Total, mean (SD)	
	Young	Old	Young	Old	Young	Old	Young	Old	Young	Old
T1: Website experience outcomes										
Website involvement	4.9 (11.0)	4.7 (1.1)	4.5 (1.0)	4.7 (1.21)	4.8 (1.1)	4.6 (1.3)	4.5 (1.1)	4.9 (1.0)	4.7 (1.1)	4.7 (1.1)
Website attractiveness	5.1[a,b] (1.0)	5.1 (1.2)	4.5 (1.1)	5.2 (1.4)	5.3[a,c] (0.9)	4.9 (1.33)	4.9 (1.2)	4.9 (1.3)	5.0 (1.1)	5.0 (1.3)
Website comprehension	6.4 (1.4)	6.2 (0.8)	6.2 (0.7)	6.0 (1.2)	6.6[b,d] (0.5)	6.0 (1.3)	6.0 (1.4)	5.8 (1.4)	6.3 (0.9)	6.0 (1.2)
Website emotional support	3.9 (1.3)	4.1 (1.2)	3.7 (1.2)	4.1 (1.7)	4.1 (1.3)	3.9 (1.4)	3.8 (1.4)	4.2 (1.5)	3.9 (1.3)	4.1 (1.4)
Self-efficacy	20.4 (2.5)	20.0 (3.4)	19.7 (3.8)	20.7 (3.0)	20.2 (2.7)	21.2 (3.00)	21.4[a,b] (3.4)	20.0 (3.1)	20.4 (3.1)	20.5 (3.2)
Anxiety	48.4 (11.2)	48.1 (11.1)	48.5 (10.1)	43.6 (11.0)	48.0 (11.2)	48.6 (8.4)	45.8 (10.9)	45.5 (10.0)	47.8 (10.8)	46.5 (10.3)
Knowledge	32.1 (12.2)	25.4 (13.4)	32.7 (13.5)	20.8 (11.6)	27.6 (11.1)	21.0 (14.1)	27.4 (15.0)	23.2 (10.4)	30.2 (13.0)	22.7[e,f] (12.6)
T2: Consultation experience outcomes										
Question asking	24.8 (21.3)	19.7 (14.9)	19.8 (13.1)	14.9 (11.9)	17.3 (11.6)	17.4 (23.2)	19.7 (17.0)	15.6 (11.2)	20.9 (17.0)	17.1 (15.9)
Anxiety	39.6[a,b] (11.0)	44.1[a,b] (12.6)	45.8 (12.8)	37.4 (10.9)	44.9 (12.0)	40.1 (12.4)	41.8 (12.1)	38.6 (9.9)	42.5 (12.0)	40.3 (11.7)
T3: Outcomes after consultation										
Knowledge from website	19.1 (12.1)	12.4 (10.8)	21.2 (13.4)	12.9 (10.9)	13.1[a,b] (11.2)	11.1 (9.4)	17.4 (15.3)	10.4 (8.3)	18.0 (13.1)	11.8[e,f] (9.9)
Information recall consultation	57.4 (11.4)	56.2 (15.2)	51.9 (13.4)	55.2 (18.8)	57.3 (16.6)	55.4 (15.5)	62.6[b,d] (15.2)	54.3 (17.1)	57.3 (14.2)	55.4 (16.3)

[a]Differs from text-only condition.

[b] Significant at *P*<.05.

[c] Significant at *P*<.10.

[d]Differs from text with video condition.

[e]Differs from younger patients.

[f]Significant at *P*<.001.

Effects on Consultation Experience Outcomes (T2)

We explored whether exposure to a mode-tailored website would influence the number of questions asked by patients during consultations (RQ1) and whether this would differ between younger and older patients (RQ2). Results showed no significant differences between condition and no interaction effects between condition and age for the number of questions asked during consultations. Additionally, we hypothesized that exposure to a mode-tailored website (vs nontailored websites) would decrease anxiety immediately after consultation (H3), with stronger effects for older patients (H4). Although we found no main effect of condition, a significant interaction between condition and age group was revealed ($F_{3,204}$=3.16; P=.03; ηp^2=0.04). However, in contrast with our expectations, older patients reported higher anxiety in the mode-tailored condition (*mean* 44.14, *SD* 12.62; P=.04) compared with the text condition (*mean* 37.36, *SD* 10.86). On the contrary, younger patients reported lower levels of anxiety in the mode-tailored condition (*mean* 39.59, *SD* 10.97; P=.046) compared with the text condition (*mean* 45.80, *SD* 12.80). Overall, the data showed no support for H3 and contrasting results for H4.

Effects on Information Recall and Knowledge After Consultation (T3)

We hypothesized that exposure to a mode-tailored (vs nontailored) website would increase patients' knowledge from the website and information recall from the consultation (H5), with differential effects for younger and older patients (H6). For both outcomes, we found no significant differences between conditions and no interaction effects. Overall, older patients acquired less knowledge from the website (*mean* 11.78, *SD* 9.88) than younger patients (*mean* 17.96, *SD* 13.12; $F_{1,194}$=12.89; P<.001; ηp^2=0.06). There were no age differences in recall from the consultation ($mean_{younger}$=57.29, *SD* 14.15; $mean_{older}$=55.35, *SD* 16.33).

What Motivation- and Ability-Related Factors Explain Knowledge and Information Recall Before and After Consultation?

Table 4 summarizes all regression models. Regarding knowledge from the website before the consultation (T1), the baseline model with individual background variables (Model 1; n=211) revealed that younger age (beta=–.23; P=.001) and higher education levels (beta=.22; P=.002) were associated with higher knowledge. Patients who received information about the clinic from other sources also reported higher knowledge (beta=.13; P=.06); however, this effect was not significant. Extending this model with website experience outcomes (Model 2) significantly improved the model (ΔR^2=0.06; P=.006; total adjusted R^2=0.17). Higher perceived website involvement (beta=.15, P=.03) and higher satisfaction with the comprehensibility of the website (beta=.15, P=.05), although the latter borderline significant, are associated with higher knowledge at T1.

Regarding information recall from the consultation (T3), no background variables were associated with information recall (Model 1, n=194). Extending the model with website experience outcomes revealed that knowledge before the consultation (beta=.22; P=.003), whether the patient had watched a video on the website (beta=.14; P=.07), and communication self-efficacy (beta=.12; P=.09) explained a significant additional proportion of variance in information recall (ΔR^2=0.12; P=.001; total adjusted R^2=0.09). The latter 2 variables were however not significantly related to information recall. No consultation experience outcomes were associated with information recall from the consultation at T3.

Regarding knowledge from the website after the consultation (T3), the baseline model with control variables (Model 1, n=185) revealed that younger age (beta=–.18; P=.02) and higher education levels (beta=.16; P=.03) were associated with higher knowledge. Extending this model with website experience outcomes significantly improved the model (ΔR^2=0.19; P<.001; total adjusted R^2=0.27). Specifically, more time spent on the website before the consultation (beta=.21; P=.002) and higher knowledge at T1 (beta=.39; P<.001) were associated with higher knowledge at T3. Age (beta=–.08; P=.23) and education level (beta=.10; P=.12) became insignificant predictors of knowledge. No consultation experience outcomes were associated with knowledge at T3.

Table 4. Regression models predicting knowledge and information recall.

Regression outcomes[a]	Website knowledge (T1; n=211)[b]				Information recall consultation (T3; n=194)[c]				Website knowledge (T3; n=185)[d]					
	Model 1		Model 2		Model 1		Model 2		Model 1		Model 2		Model 3	
	Beta	*P* value	Beta	*P* value	Beta	*P* value	Beta	*P* value	Beta	*P* value	Beta	*P* value	Beta	*P* value
Individual background characteristics														
Age (years)	-0.23	.001	-0.23	.001	-0.08	.26	-0.04	.59	-0.18	.02	-0.08	.23	-0.08	.24
High education level	0.22	.002	0.22	.002	-0.02	.84	-0.05	.51	0.16	.03	0.1	.12	0.1	.14
Internet use	-0.02	.75	-0.02	.72	—[e]	—	—	—	0.09	.22	0.05	.48	0.05	.49
Coping style	0.11	.11	0.04	.61	—	—	—	—	—	—	—	—	—	—
Additional information received	0.13	.05	0.1	.11	—	—	—	—	—	—	—	—	—	—
Quality of life	—	—	—	—	—	—	—	—	-0.13	.08	-0.1	.15	-0.1	.16
Website experience characteristics														
Website involvement	—	—	0.15	.03	—	—	—	—	—	—	—	—	—	—
Website attractiveness	—	—	-0.01	.87	—	—	0.08	.32	—	—	—	—	—	—
Website comprehension	—	—	0.15	.05	—	—	-0.02	.78	—	—	—	—	—	—
Watched a video	—	—	0.08	.22	—	—	0.14	.07	—	—	—	—	—	—
Knowledge (T1)	—	—	—	—	—	—	0.22	.003	—	—	0.39	.001	0.39	.001
Website emotional support	—	—	—	—	—	—	0.07	.36	—	—	—	—	—	—
Communication self-efficacy	—	—	—	—	—	—	0.12	.09	—	—	—	—	—	—
Time on website	—	—	—	—	—	—	0.07	.36	—	—	0.21	.02	0.21	.002
Anxiety (T1)	—	—	—	—	—	—	—	—	—	—	0.04	.52	0.04	.52
Consultation experience characteristics														
Question asking	—	—	—	—	—	—	—	—	—	—	—	—	0	.96

[a]Only variables marginally significant ($P<.10$) that correlated with the predicted outcome variable were included. As no consultation characteristics correlated with information recall from the consultation (T3), only 2 models were predicted. Model 1 shows a simple linear regression model assessing the relationship between control variables and knowledge/information recall. Website experience characteristics were added to Model 2. Consultation experience was included in Model 3. We report the models without controlling for health literacy owing to missing data. Repeating the analyses with health literacy in the models did not change results, although health literacy significantly related to website knowledge (T3). R^2 indicates the adjusted explained variance of the model; ΔR^2 shows the change in R^2 by adding predictors in Model 2 and 3; significant ΔF shows whether the difference in the F value for model expansion is significant.

[b]Adjusted R^2 Model 1=0.12, Model 2=0.17). Adding website experience characteristics to Model 2 improved the model (ΔR^2=0.06, P=.006).

[c]Adjusted R^2 Model 1=0.00, Model 2=0.09). Adding website experience characteristics to Model 2 improved the model (ΔR^2=0.12, P=.001).

[d]Adjusted R^2 Model 1=0.09, Model 2=0.27, Model 3=0.27). Adding website experience characteristics to Model 2 improved the model (ΔR^2=0.19, $P<.001$). Addition consultation experience characteristics to Model 3 did not improve the model (ΔR^2=0.00, P=.96).

[e]Not applicable.

Discussion

Principal Findings

This RCT tested the effectiveness of a mode-tailored preparatory website (ie, by self-selecting text, images, and/or videos) versus nontailored websites (ie, with either text only, text with images, or text with videos) in a clinical population of older (≥65 years) and younger (<65) patients visiting a fast-track clinic for diagnosis and treatment planning for colorectal, esophageal, or stomach cancer. The main research question was whether mode tailoring is more effective than nontailored information on patient-reported outcomes before, during, and after consultation. Moreover, we investigated whether older patients benefited proportionally more from mode-tailored information than younger patients. To advance theoretical models on information processing and the interplay between online information provision and offline patient-provider communication, we

additionally explored which *website experience outcomes* and *consultation experience outcomes* contributed to patients' knowledge acquisition from online health information and information recall from consultations. Following is a review of the study results in light of the unexpected findings, discuss the implications for theory, and suggest directions for future research.

Review of Findings

The first main and unexpected finding of this study was that younger patients were more satisfied with the mode-tailored website (vs text only), whereas this was not the case for older patients. Moreover, younger patients who viewed the mode-tailored website reported lower anxiety levels immediately after consultation (vs text only). In contrast with our hypothesis, older patients reported higher anxiety in the mode-tailored condition (vs text only). Posthoc analyses revealed that younger patients also showed greater anxiety *reduction* from pre- to postconsultation, after viewing the mode-tailored website (vs the nontailored websites), whereas the anxiety levels of older patients in the mode-tailored condition remained the same. Both younger and older patients had equal anxiety levels before consultation. On the basis of the socioemotional selectivity theory, we expected that older patients would perceive more emotional gratification if they had the option to select information presented in visual/audiovisual modes, thereby limiting their anxiety. Alternatively, we now discuss a different explanation for our findings, from an uncertainty management theory perspective combined with the socioemotional selectivity theory. Generally speaking, younger adults are less tolerant to uncertainty than older adults [55] and, therefore, more likely to seek information as a strategy to reduce uncertainty [56,57]. It could be that, in our study, younger patients were more intolerant to the uncertainty that came with their cancer diagnosis and, therefore, exhausted all their information sources (ie, different information modes) to reduce uncertainty and, thereby, their anxiety. Older adults, on the contrary, are generally better at tolerating and managing uncertainty [55] and could be less urged to reduce it by means of information. Information might even have reversed effects and increase anxiety in this group, especially among *older-old* patients with cancer (≥70 years), who more often prefer to leave information disclosure up to the health care provider [58]. Even though younger and older patients did not differ in information seeking/avoidance in this study, age-related differences in uncertainty tolerance might explain why providing information in tailored multiple modes particularly benefited younger patients, while it had no effect on older patients. The socioemotional selectivity theory also presents an explanation for this finding. Namely, younger adults pertain more to knowledge-related goals to prepare for future events, whereas older adults attach greater importance to emotionally meaningful goals [33]. It is possible that being able to view information in different (visual) modalities in a tailored manner accommodated to younger patients' information needs (ie, knowledge-related goals) and supported them in lowering their anxiety, which was less so for older patients. Future research is warranted to understand the role of uncertainty intolerance and information seeking/avoidance, as well as knowledge- versus emotional-related goals, in elucidating whether and how online tailored health information might accommodate the needs of and benefit younger and older patients with cancer.

Overall, we did not find that mode tailoring proportionally benefitted older patients more than younger patients. Possibly, older patients had more difficulty using the mode tailoring functionality, whereas this was more intuitive for younger patients, because of differences in internet experience (see Table 1). More long-term use and experience with the tailoring tool could make the mode-tailored website become beneficial for older patients as well. Alternatively, in our study sample, many patients were in their 60s (41.8%). Despite the significant age difference of younger and older patients, many patients were aged around the cut-off of 65 years. This could explain why the age groups did not differ on age-related background variables and why no age differences were found in outcome variables. To illustrate, older patients in this study were not frailer than younger patients. In fact, the mean frailty score was 2.46 (*SD* 1.98), with less than a quarter of patients (23.6%) reporting a score higher than 4, a cut-off used to identify patients as (moderately) frail [49]. Recent work found that age-related factors (eg, frailty, health literacy, and future time perspective) are more predictive of information recall from cancer websites than chronological age [40]. Moreover, older adults are a highly heterogeneous group in fundamental domains such as biological, cognitive, and personality characteristics [59], which could influence how online health information is used, processed, and evaluated [60]. Therefore, when investigating website use behaviors and intervention effects on patient outcomes, it might be meaningful to consider age-related variables as moderators of effects.

The second main finding of this study is that certain website experience outcomes (eg, website involvement and time spent on the website) increased patients' knowledge before and after the consultation. Higher previsit knowledge in turn supports patients in recalling information from the consultation. This suggests that offering information to help patients prepare for their hospital visit can improve past knowledge, which is important for how they process information during consultations, and remember this afterward. This is an important outcome, as knowledge is one of the key prerequisites for patients to be able to be involved in making treatment decisions and manage their illness [61]. Moreover, the website was widely used by patients across all website conditions (*mean* 34 min), and this was not affected by age, gender, or education level. The majority of patients even used the website multiple times before their hospital visit. This underlines a desire for information before a hospital visit, perhaps even more so in emotionally charged contexts such as the diagnosis and treatment planning phase. What makes these findings particularly noteworthy is that in this study, we did not find a relationship between anxiety during consultation and information recall from the consultation. In a previous observational field study among the same patient population, at the same outpatient clinic, but without a preparatory website intervention, anxiety during consultation negatively predicted information recall from consultations [6]. Interestingly, in this RCT, anxiety was not found to be a barrier for information recall. A comparison between the 2 samples

reveals that patients in both studies reported equally high anxiety levels. This raises the question: Could it be that offering patients' preparatory online information before their hospital visit helped them to attend to and process information from the consultation despite their anxiety levels? Previous studies showed that highly anxious patients with cancer have higher information needs [62] and that patients with fulfilled information needs are less anxious [63]. These findings, together with the results from our previous observational study and this RCT, suggest that knowledge (ie, by means of online health information) may play an important role in patients' anxiety management overall, especially for younger patients. However, to answer this question with more accuracy, future research is warranted to understand the added value of offering online preparatory information (vs no information) on patients' fulfillment of information needs, knowledge/information recall, and the anxiety-recall relationship.

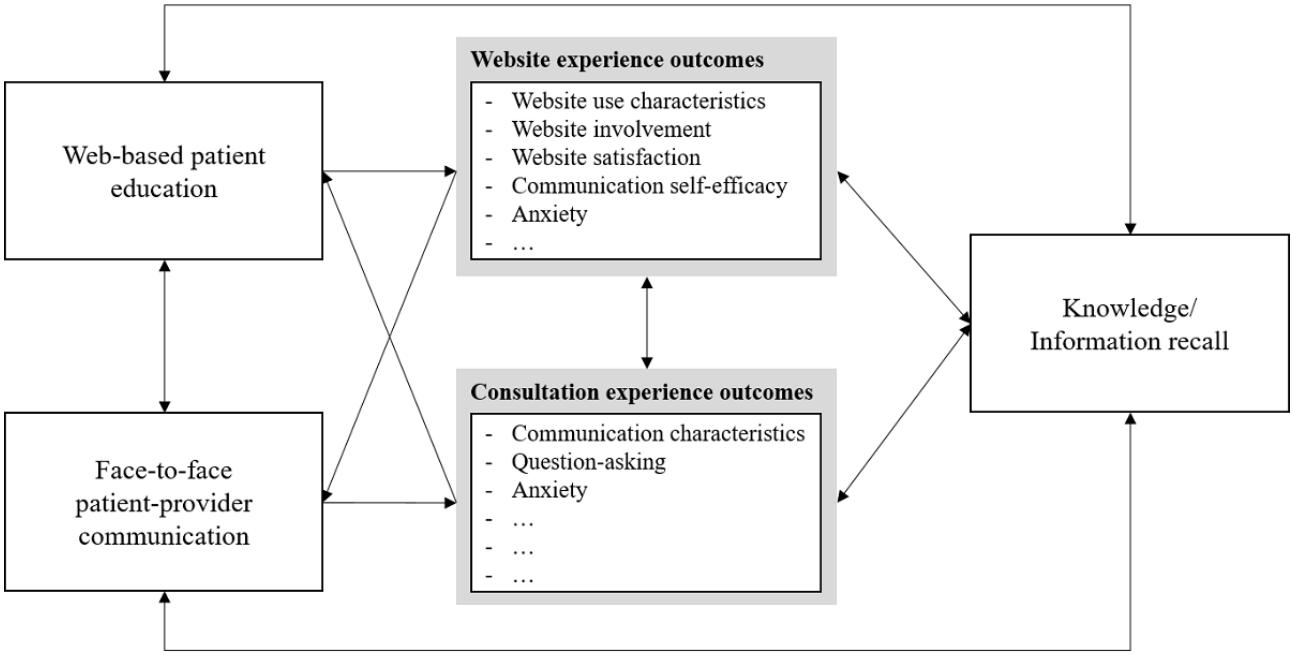
Although no differences were found on main outcome variables, our data showed that patients spent more time on the mode-tailored website than on the nontailored website versions. This difference was not significant, but it suggests that mode tailoring may trigger patients to attend to the website information longer, which in this study proved to be important for knowledge acquisition from the website and information recall from consultations. The results align with earlier experimental findings that mode tailoring online information can increase attention to website information and consequently enhance information recall [17]. Although it is possible that the extra time spent on the mode-tailored website was because patients needed more time to figure out how the mode tailoring tool worked, it is more likely that patients spent this time viewing the website content. Namely, the time range that patients needed to select their first mode does not weigh up to the extra time patients spent on the mode-tailored website (vs the nontailored websites). Moreover, posthoc analyses revealed that viewing the mode-tailored website required comparable levels of cognitive effort as the nontailored versions for both younger and older patients (mean 2.65, *SD* 1.05; range 1-7 with higher scores indicating higher cognitive load, $F_{3,229}$=0.23; *P*=.88; ηp^2=0.00). In addition, even though no clear differences were found between conditions on the hypothesized outcome variables, the mode-tailored website kept patients online for a slightly longer period of time, which is likely to be important for knowledge gain. Together, the results suggest that offering online information to patients as a preparation tool might benefit patient outcomes.

Theoretical Implications

This study aimed to gain insight into which *website experience outcomes* and *consultation experience outcomes* explain the benefits of preparatory online information on knowledge acquisition from websites and information recall from consultations in the cancer communication context. As such, the results of this study are helpful in refining existing theoretical frameworks that conceptualize the interplay between mediated health communication and offline patient-provider communication in an eHealth era. At present, new information technologies open up a wide range of possibilities for patients to obtain health information outside of the consultation room. Scholars have conceptualized models of patient-provider communication that consider the role of mediated health communication to explain how communication can affect health outcomes via different pathways, such as by increasing patient knowledge, enhancing patients' ability to manage emotions, and enhancing patient empowerment and agency [64,65]. Although such overviews are a useful starting point, our findings help further specify which specific pathways, considering both online information provision and offline patient-provider communication, are key in improving patient health outcomes. Specifically, we identified variables related to attention (ie, time spent online) and website involvement as important *motivational website experienceoutcomes* that contribute to processing of online information, thereby increasing patients' knowledge. Although not one of the *consultation experience outcomes* (ie, question asking and anxiety) was related to information recall from consultations, we identified increased knowledge levels before consultation as an *ability-related website experience outcome* that benefitted patients' recall of information from the consultations. These results also inform future research relying on theories such as the ELM and LC4MP to understand which specific motivation- and ability-related factors play a role in how online cancer information is processed and how this may impact patient outcomes. This is important, as a critique on the ELM has been that its mediating variables are not clearly defined, leaving room for further specification of which variables contribute to motivation and ability to process information in different situational contexts [66]. We conclude with an integrated model visualizing how the use of online health information and patient-provider communication may lead to website experience outcomes and consultation experience outcomes, which may reinforce each other and ultimately explain knowledge and information recall (Figure 4). The process should be seen as a cycle, where each outcome can in turn influence the use of online health information and shape interactions between patients and providers. We note that this model is merely a *starting point*, and future research should complement this with relevant theoretical concepts that were unconsidered in our study.

This study also adds to the existing literature on computer-tailored health communication. Tailored online health information tools, typically providing recipients with *content* adapted to their individual characteristics, needs, and/or preferences have shown to be more effective than nontailored information on a wide range of patient outcomes, however typically with small effects [67-71]. Such tailored communication interventions may have provided personally relevant *content* but have possibly overlooked individual preferences for *how* this information should be presented. Moreover, the way in which content is processed highly depends on how this information is delivered [72]. Even though effects were small, this study suggests that tailoring the mode of information presentation to individual preferences, abilities, and/or learning styles could enhance the effectiveness of online health information interventions. Future research combining different tailoring strategies, such as content tailoring and mode tailoring, is warranted to tell whether effect sizes of tailored health communication interventions can be improved.

Figure 4. Integrated model of website experience outcomes and consultation experience outcomes explaining knowledge/information recall.



Strengths, Limitations, and Future Research Directions

This study is the first to translate mode tailoring research to a clinical population of patients with cancer. Major strengths are its longitudinal character (ie, following patients from before to during and after their hospital visit) and the combination of observational data (ie, website tracking data and video observations) and patient-reported outcomes. We were able to include over 240 patients in our trial, which is highly unique given the emotional burden that newly diagnosed patients with cancer experience and including patients at this stage is challenging. About half (53.3%, 276/517) of patients declined participation. Although there was no age and gender difference between participating and nonparticipating patients, it is possible that those with varying education or health literacy levels were more or less likely to participate. Other RCTs employing educational interventions in cancer care report inclusion rates approximately between 25% and 60% [73-75]. Given that our inclusion rate was 47%, we believe that our carefully crafted inclusion protocol made it possible to reach the desired sample size in a difficult-to-reach population (see the *Methods* section).

Previous tailoring studies examining modes of information presentation have been useful in identifying which message features yield effects on outcomes and in unraveling the theoretical mechanisms explaining these effects [17,18,76]. However, these findings must be translated to different clinical contexts to establish whether such interventions have added value in different real-life settings. Several null findings in this study, concerning website experience outcomes (eg, communication self-efficacy) and consultation-related outcomes (eg, question asking), cannot be compared with the previous research. Below we discuss possible reasons why mode tailoring did not produce similar effects in this study's clinical setting compared with previous experimental studies [17,18].

The first explanation is the age discrepancy between our previous experimental work and this RCT. *Younger* patients, that is, those aged younger than 65 years, were for the greater part not represented in the experimental studies as they included younger adults between the ages of 25 and 45 years [17,18]. Moreover, as participants in the experimental studies were recruited via an online research panel, these older participants were likely to be healthier and have more internet experience than clinical older patients in this RCT.

The second explanation is the high topic involvement of patients and therefore high personal relevance of the website content, irrespective of how it was presented. The ELM suggests that information is more likely to be processed deeply when a person's interest for certain information is high, resulting in greater effects on outcomes (eg, information recall) [20]. Our sample consisted of newly diagnosed patients who were directed to a website about the specific clinic they were referred to, with information about their specific condition. Additionally, information avoiders might be underrepresented in our sample, as they might be less inclined to participate in our study. This might explain why the website was well used across all conditions (with only 9 patients not viewing the website) and why the perceived relevance of information did not differ between conditions. Consequently, it is possible that information was processed equally well from all website versions, revealing no differences in outcomes between conditions. Relatedly, it could be that some patients in the nontailored conditions received information in a way that coincidentally matched their preferences, attenuating effects of the mode-tailored website on outcomes. A previous tailoring study showed that when standardized information (by chance) corresponded to individual information needs, this was just as effective as tailored information [77]. Alternative explanations could be that the life-threatening nature of the disease (cancer), the emotionally charged moment (diagnosis), or a combination of these two elevated the perceived relevance and, consequently, website use. In this study, patients received the information while awaiting a final diagnosis and treatment plan, which is a phase in which information needs are the highest [26]. Future research

could investigate whether mode tailoring has added value for clinical patients with a less life-threatening disease (eg, asthma, diabetes, and hypertension).

The third explanation is the uncontrolled setting of a field trial. In experimental studies, participants are exposed to stimulus materials, such as websites, and asked about these materials immediately afterward. In our field study, patients received a link to the website and the baseline questionnaire several days before their hospital visit. This allowed for patients to fill out the baseline questionnaire assessing website experience variables (eg, website satisfaction, anxiety, and communications self-efficacy) on different days before their visit. Moreover, patients varied in how often and how long they consulted the website. The variability concerning *when* their answers were recorded and *how* they used the website could have diluted the observable effects (eg, knowledge scores), if present. Although it would have been ideal to standardize study procedures even more, this is difficult, and perhaps unfruitful, in a clinical field study with patients with cancer.

In a similar vein, our sample consisted of a heterogeneous group of patients dealing with different cancer types (ie, colorectal, esophageal, or stomach cancer) with varying health trajectories before their appointment at the outpatient clinic. Although we imposed a strict randomization procedure that showed that patients did not differ on patient background characteristics, the sample heterogeneity could have attenuated the observable effects. We managed to recruit a relatively large sample for a difficult-to-reach clinical population, but it remains possible that we were unable to detect some of the hypothesized effects owing to a lack of statistical power. As a solution, qualitative approaches (eg, observations of or interviews with patients who have used the website intervention) might give a more meaningful and in-depth analysis of how the website was used, by whom, and whether this had added value for patients. This might be especially true for health communication interventions where small effect sizes are expected, and it remains difficult to obtain large, relatively homogeneous samples.

This study design did not include a no-information control group. A no-information control group would be useful to examine whether offering online preparatory information in general would have added value, but we considered that it would add little insight into mode tailoring effects. Furthermore, including a no-information control group would significantly reduce the number of patients in each condition given that data collection was limited to a maximum period of 3 years, further limiting statistical power to detect group differences. Although a no-information control group was not included, the results show that all the variables that predicted knowledge and recall were associated with the website (website involvement, website comprehension, and time spent on the website) and, therefore, imply that additional information sources surrounding the consultation can benefit patients.

In conclusion, we believe that there are many fruitful directions for future research. In addition to the suggestions we have already made, future research could for instance explore the relative importance of online information provision compared with offline, more traditional methods of information provision in different contexts and patient populations. To optimize information provision to all patients, researchers should continue to explore the added benefit of providing online preparatory information to patients (eg, in the form of hospital websites and patient portals), how specific features of the internet (eg, modality and interactivity) can be used to tailor information to patients, and whether different tailoring strategies (eg, content, mode, and cultural tailoring) are effective for patients regarding different types of patient outcomes (eg, evaluative, cognitive, psychosocial and behavioral) and in different patient populations (eg, high emotionally charged settings and non–life-threatening chronic diseases).

Conclusions

This RCT showed that higher use of online health information to prepare for consultations benefits patients' knowledge levels before a hospital visit. Higher knowledge, in turn, facilitates information processing and results in better information recall from medical consultations and knowledge acquisition from online information after their visit. Moreover, viewing online health information in a tailored presentation mode (ie, textual, visual, and/or audiovisual) increased younger patients' satisfaction with the health website and reduced their anxiety after consultation, but not for older patients. The results are important in refining existing theoretical frameworks of patient-provider communication in an eHealth era.

Acknowledgments

The authors thank all patients who participated in this study, the health care providers at the GIOCA for facilitating data collection, and all research assistants who helped collect the data. This study was funded by the Dutch Cancer Society (KWF Kankerbestrijding) under grant number UvA 2014-6700. KWF had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Conflicts of Interest

None declared.

References

1. Pilleron S, Sarfati D, Janssen-Heijnen M, Vignat J, Ferlay J, Bray F, et al. Global cancer incidence in older adults, 2012 and 2035: a population-based study. Int J Cancer 2019 Jan 1;144(1):49-58. [doi: 10.1002/ijc.31664] [Medline: 29978474]

2. Sparks L, Turner MM. The impact of cognitive and emotive communication barriers on older adult message processing of cancer-related health information: new directions for research. In: Sparks L, O'Hair H, Kreps G, editors. Cancer, Communication and Aging. New York: Hampton Press; 2008:17-47.
3. Butow PN, Dunn SM, Tattersall MH, Jones QJ. Patient participation in the cancer consultation: evaluation of a question prompt sheet. Ann Oncol 1994 Mar;5(3):199-204. [doi: 10.1093/oxfordjournals.annonc.a058793] [Medline: 8186168]
4. Maliski SL, Kwan L, Krupski T, Fink A, Orecklin JR, Litwin MS. Confidence in the ability to communicate with physicians among low-income patients with prostate cancer. Urology 2004 Aug;64(2):329-334. [doi: 10.1016/j.urology.2004.03.042] [Medline: 15302489]
5. Jansen J, Butow PN, van Weert JC, van Dulmen S, Devine RJ, Heeren TJ, et al. Does age really matter? Recall of information presented to newly referred patients with cancer. J Clin Oncol 2008 Nov 20;26(33):5450-5457. [doi: 10.1200/JCO.2007.15.2322] [Medline: 18936478]
6. Nguyen MH, Smets EM, Bol N, Bronner MB, Tytgat KM, Loos EF, et al. Fear and forget: how anxiety impacts information recall in newly diagnosed cancer patients visiting a fast-track clinic. Acta Oncol 2019 Feb;58(2):182-188. [doi: 10.1080/0284186X.2018.1512156] [Medline: 30264632]
7. Medlock S, Eslami S, Askari M, Arts DL, Sent D, de Rooij SE, et al. Health information-seeking behavior of seniors who use the internet: a survey. J Med Internet Res 2015 Jan 8;17(1):e10 [FREE Full text] [doi: 10.2196/jmir.3749] [Medline: 25574815]
8. Luck A, Pearson S, Maddern G, Hewett P. Effects of video information on precolonoscopy anxiety and knowledge: a randomised trial. Lancet 1999 Dec 11;354(9195):2032-2035. [doi: 10.1016/s0140-6736(98)10495-6] [Medline: 10636368]
9. McMullan M. Patients using the internet to obtain health information: how this affects the patient-health professional relationship. Patient Educ Couns 2006 Oct;63(1-2):24-28. [doi: 10.1016/j.pec.2005.10.006] [Medline: 16406474]
10. Brandes K, Linn AJ, Butow PN, van Weert JC. The characteristics and effectiveness of question prompt list interventions in oncology: a systematic review of the literature. Psychooncology 2015 Mar;24(3):245-252. [doi: 10.1002/pon.3637] [Medline: 25082386]
11. van Weert JC, Jansen J, Spreeuwenberg PM, van Dulmen S, Bensing JM. Effects of communication skills training and a question prompt sheet to improve communication with older cancer patients: a randomized controlled trial. Crit Rev Oncol Hematol 2011 Oct;80(1):145-159. [doi: 10.1016/j.critrevonc.2010.10.010] [Medline: 21075644]
12. Bolle S, Romijn G, Smets EM, Loos EF, Kunneman M, van Weert JC. Older cancer patients' user experiences with web-based health information tools: a think-aloud study. J Med Internet Res 2016 Jul 25;18(7):e208 [FREE Full text] [doi: 10.2196/jmir.5618] [Medline: 27457709]
13. Friemel TN. The digital divide has grown old: determinants of a digital divide among seniors. New Media Soc 2014 Jun 12;18(2):313-331. [doi: 10.1177/1461444814538648]
14. Loos EF, Romano-Bergstrom J. Older adults. In: Romano-Bergstrom J, Schall AJ, editors. Eye Tracking in User Experience Design. Amsterdam: Elsevier; 2014:313-329.
15. Wright P, Soroka A, Belt S, Pham D, Dimov S, DeRoure D, et al. Modality preference and performance when seniors consult online information. Gerontechnology 2008;7(3):293-304. [doi: 10.4017/gt.2008.07.03.004.00]
16. Smit ES, Linn AJ, van Weert JC. Taking online computer-tailoring forward: the potential of tailoring the message frame and delivery mode of online health behaviour change interventions. Eur Health Psychol 2015;17(1):25-31 [FREE Full text]
17. Nguyen MH, van Weert JC, Bol N, Loos EF, Tytgat KM, van de Ven AW, et al. Tailoring the mode of information presentation: effects on younger and older adults' attention and recall of online information. Hum Commun Res 2016 Oct 21;43(1):102-126. [doi: 10.1111/hcre.12097]
18. Nguyen MH, Smets EM, Bol N, Loos EF, van Weert JC. How tailoring the mode of information presentation influences younger and older adults' satisfaction with health websites. J Health Commun 2018;23(2):170-180. [doi: 10.1080/10810730.2017.1421729] [Medline: 29345531]
19. Lang A. Using the limited capacity model of motivated mediated message processing to design effective cancer communication messages. J Commun 2006 Aug;56(s1):S57-S80 [FREE Full text] [doi: 10.1111/j.1460-2466.2006.00283.x]
20. Petty R, Barden J, Wheeler S. The elaboration likelihood model of persuasion: health promotions that yield sustained behavior change. In: DiClemente RJ, Crosby RA, Kegler MC, editors. Emerging Theories in Health Promotion Practice and Research. San Fransisco: Jossey-Bass; 2002:71-99.
21. Bodie GD, Dutta MJ. Understanding health literacy for strategic health marketing: ehealth literacy, health disparities, and the digital divide. Health Mark Q 2008;25(1-2):175-203. [doi: 10.1080/07359680802126301] [Medline: 18935884]
22. Rideout V, Neuman T, Kitchman M, Brodie M. Kaiser Family Foundation. 2005. e-Health and the Elderly: How Seniors Use the Internet for Health Information URL: https://www.kff.org/wp-content/uploads/2013/01/e-health-and-the-elderly-how-seniors-use-the-internet-for-health-information-key-findings-from-a-national-survey-of-older-americans-survey-report.pdf [accessed 2017-03-30] [WebCite Cache ID 6pLm4uZzn]
23. Becker SA. A study of web usability for older adults seeking online health resources. ACM Trans Comput-Hum Interact 2004 Dec 1;11(4):387-406. [doi: 10.1145/1035575.1035578]

24. Bronner MB, Nguyen MH, Smets EM, van de Ven AW, van Weert JC. Anxiety during cancer diagnosis: examining the influence of monitoring coping style and treatment plan. Psychooncology 2018 Feb;27(2):661-667. [doi: 10.1002/pon.4560] [Medline: 28976610]
25. Rees CE, Bath PA. Information-seeking behaviors of women with breast cancer. Oncol Nurs Forum 2001 Jun;28(5):899-907. [Medline: 11421149]
26. Rutten LJ, Arora NK, Bakos AD, Aziz N, Rowland J. Information needs and sources of information among cancer patients: a systematic review of research (1980-2003). Patient Educ Couns 2005 Jun;57(3):250-261. [doi: 10.1016/j.pec.2004.06.006] [Medline: 15893206]
27. Nguyen MH, Bol N, van Weert JC, Loos EF, Tytgat KM, Geijsen D, et al. Optimising ehealth tools for older patients: collaborative redesign of a hospital website. Eur J Cancer Care (Engl) 2019 Jan;28(1):e12882 [FREE Full text] [doi: 10.1111/ecc.12882] [Medline: 30015998]
28. Albada A, van Dulmen S, Lindhout D, Bensing JM, Ausems MG. A pre-visit tailored website enhances counselees' realistic expectations and knowledge and fulfils information needs for breast cancer genetic counselling. Fam Cancer 2012 Mar;11(1):85-95. [doi: 10.1007/s10689-011-9479-1] [Medline: 21901499]
29. Jerant A, Sohler N, Fiscella K, Franks B, Franks P. Tailored interactive multimedia computer programs to reduce health disparities: opportunities and challenges. Patient Educ Couns 2011 Nov;85(2):323-330 [FREE Full text] [doi: 10.1016/j.pec.2010.11.012] [Medline: 21146950]
30. Street Jr RL, Slee C, Kalauokalani DK, Dean DE, Tancredi DJ, Kravitz RL. Improving physician-patient communication about cancer pain with a tailored education-coaching intervention. Patient Educ Couns 2010 Jul;80(1):42-47 [FREE Full text] [doi: 10.1016/j.pec.2009.10.009] [Medline: 19962845]
31. Linn AJ, van Weert JC, Smit EG, Perry K, van Dijk L. 1+1=3? The systematic development of a theoretical and evidence-based tailored multimedia intervention to improve medication adherence. Patient Educ Couns 2013 Dec;93(3):381-388. [doi: 10.1016/j.pec.2013.03.009] [Medline: 23579041]
32. Iverson SA, Howard KB, Penney BK. Impact of internet use on health-related behaviors and the patient-physician relationship: a survey-based study and review. J Am Osteopath Assoc 2008 Dec;108(12):699-711. [doi: 10.7556/jaoa.2008.108.12.699] [Medline: 19075034]
33. Löckenhoff CE, Carstensen LL. Socioemotional selectivity theory, aging, and health: the increasingly delicate balance between regulating emotions and making tough choices. J Pers 2004 Dec;72(6):1395-1424. [doi: 10.1111/j.1467-6494.2004.00301.x] [Medline: 15509287]
34. Ogozalek VZ. A Comparison of the Use of Text and Multimedia Interfaces to Provide Information to the Elderly. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 1994 Presented at: CHI'94; April 24-28, 1994; Boston, Massachusetts, USA p. 65-71. [doi: 10.1145/191666.191700]
35. Bol N, Smets EM, Eddes EH, de Haes JC, Loos EF, van Weert JC. Illustrations enhance older colorectal cancer patients' website satisfaction and recall of online cancer information. Eur J Cancer Care (Engl) 2015 Mar;24(2):213-223. [doi: 10.1111/ecc.12283] [Medline: 25615269]
36. Bol N, Smets EM, Rutgers MM, Burgers JA, de Haes HC, Loos EF, et al. Do videos improve website satisfaction and recall of online cancer-related information in older lung cancer patients? Patient Educ Couns 2013 Sep;92(3):404-412. [doi: 10.1016/j.pec.2013.06.004] [Medline: 23820196]
37. Soroka AJ, Wright P, Belt S, Pham DT, Dimov SS, de Roure D, et al. User Choices for Modalities of Instructional Information. In: Proceedings of the 4th IEEE International Conference on Industrial Informatics. 2006 Presented at: INDIN'06; August 16-18, 2006; Singapore, Singapore p. 411-416. [doi: 10.1109/INDIN.2006.275835]
38. Kessels RP. Patients' memory for medical information. J R Soc Med 2003 May;96(5):219-222 [FREE Full text] [doi: 10.1258/jrsm.96.5.219] [Medline: 12724430]
39. Posma ER, van Weert JC, Jansen J, Bensing JM. Older cancer patients' information and support needs surrounding treatment: an evaluation through the eyes of patients, relatives and professionals. BMC Nurs 2009 Jan 19;8:1 [FREE Full text] [doi: 10.1186/1472-6955-8-1] [Medline: 19152675]
40. Bol N, Smets EM, Burgers JA, Samii SM, de Haes HC, van Weert JC. Older patients' recall of online cancer information: do ability and motivation matter more than chronological age? J Health Commun 2018;23(1):9-19. [doi: 10.1080/10810730.2017.1394400] [Medline: 29227736]
41. Basta YL, Baur O, van Dieren DS, Klinkenbijl JH, Fockens P, Tytgat KM. Is there a benefit of multidisciplinary cancer team meetings for patients with gastrointestinal malignancies? Ann Surg Oncol 2016 Aug;23(8):2430-2437 [FREE Full text] [doi: 10.1245/s10434-016-5178-3] [Medline: 27002814]
42. Dutta-Bergman MJ. The impact of completeness and web use motivation on the credibility of e-health information. J Commun 2004;54(2):253-269. [doi: 10.1111/j.1460-2466.2004.tb02627.x]
43. Bol N, van Weert JC, de Haes HC, Loos EF, de Heer S, Sikkel D, et al. Using cognitive and affective illustrations to enhance older adults' website satisfaction and recall of online cancer-related information. Health Commun 2014;29(7):678-688. [doi: 10.1080/10410236.2013.771560] [Medline: 24160560]

44. ten Klooster KP, Oostveen JC, Zandbelt LC, Taal E, Drossaert CH, Harmsen EJ, et al. Further validation of the 5-item perceived efficacy in patient-physician interactions (PEPPI-5) scale in patients with osteoarthritis. Patient Educ Couns 2012 Apr;87(1):125-130. [doi: 10.1016/j.pec.2011.07.017] [Medline: 21889864]
45. Marteau TM, Bekker H. The development of a six-item short-form of the state scale of the Spielberger state-trait anxiety inventory (STAI). Br J Clin Psychol 1992 Sep;31(Pt 3):301-306. [doi: 10.1111/j.2044-8260.1992.tb00997.x] [Medline: 1393159]
46. Millar K, Jelicic M, Bonke B, Asbury AJ. Assessment of preoperative anxiety: comparison of measures in patients awaiting surgery for breast cancer. Br J Anaesth 1995 Feb;74(2):180-183 [FREE Full text] [doi: 10.1093/bja/74.2.180] [Medline: 7696068]
47. Jansen J, van Weert JC, van der Meulen N, van Dulmen S, Heeren T, Bensing JM. Recall in older cancer patients: measuring memory for medical information. Gerontologist 2008 Apr;48(2):149-157. [doi: 10.1093/geront/48.2.149] [Medline: 18483427]
48. Zandbelt LC, Smets EM, Oort FJ, Godfried MH, de Haes HC. Medical specialists' patient-centered communication and patient-reported outcomes. Med Care 2007 Apr;45(4):330-339. [doi: 10.1097/01.mlr.0000250482.07970.5f] [Medline: 17496717]
49. Maat HP, Essink-Bot ML, Leenaars KE, Fransen MP. A short assessment of health literacy (SAHL) in the Netherlands. BMC Public Health 2014 Sep 23;14:990 [FREE Full text] [doi: 10.1186/1471-2458-14-990] [Medline: 25246170]
50. Steverink N, Slaets J, Schuurmans H, van Lis M. Measuring frailty: development and testing of the Groningen frailty indicator (GFI). Gerontologist 2001;41(1):236-237 [FREE Full text]
51. Aaronson NK, Ahmedzai S, Bergman B, Bullinger M, Cull A, Duez NJ, et al. The European Organization for Research and Treatment of Cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology. J Natl Cancer Inst 1993 Mar 3;85(5):365-376. [doi: 10.1093/jnci/85.5.365] [Medline: 8433390]
52. Miller SM. Monitoring and blunting: validation of a questionnaire to assess styles of information seeking under threat. J Pers Soc Psychol 1987 Feb;52(2):345-353. [doi: 10.1037/0022-3514.52] [Medline: 3559895]
53. Thomas R, Kaminski E, Stanton E, Williams M. Measuring information strategies in oncology - developing an information satisfaction questionnaire. Eur J Cancer Care (Engl) 2004 Mar;13(1):65-70. [doi: 10.1111/j.1365-2354.2004.00445.x] [Medline: 14961777]
54. van Weert JC, Jansen J, de Bruijn GJ, Noordman J, van Dulmen S, Bensing JM. QUOTEchemo: a patient-centred instrument to measure quality of communication preceding chemotherapy treatment through the patient's eyes. Eur J Cancer 2009 Nov;45(17):2967-2976. [doi: 10.1016/j.ejca.2009.06.001] [Medline: 19615889]
55. Basevitz P, Pushkar D, Chaikelson J, Conway M, Dalton C. Age-related differences in worry and related processes. Int J Aging Hum Dev 2008;66(4):283-305. [doi: 10.2190/AG.66.4.b] [Medline: 18507331]
56. Carcioppolo N, Yang F, Yang Q. Reducing, maintaining, or escalating uncertainty? The development and validation of four uncertainty preference scales related to cancer information seeking and avoidance. J Health Commun 2016 Sep;21(9):979-988. [doi: 10.1080/10810730.2016.1184357] [Medline: 27450905]
57. Miller LE. Uncertainty management and information seeking in cancer survivorship. Health Commun 2014;29(3):233-243. [doi: 10.1080/10410236.2012.739949] [Medline: 23682727]
58. Jenkins V, Fallowfield L, Saul J. Information needs of patients with cancer: results from a large study in UK cancer centres. Br J Cancer 2001 Jan 5;84(1):48-51 [FREE Full text] [doi: 10.1054/bjoc.2000.1573] [Medline: 11139312]
59. Stone ME, Lin J, Dannefer D, Kelley-Moore JA. The continued eclipse of heterogeneity in gerontological research. J Gerontol B Psychol Sci Soc Sci 2017 Jan;72(1):162-167 [FREE Full text] [doi: 10.1093/geronb/gbv068] [Medline: 26834089]
60. van der Goot M. Older adults' mediated communication: current perspectives among communication scholars. In: Nussbaum J, editor. Communication Across the Lifespan. New York: Peter Lang; 2016:79-89.
61. Elwyn G, Frosch D, Thomson R, Joseph-Williams N, Lloyd A, Kinnersley P, et al. Shared decision making: a model for clinical practice. J Gen Intern Med 2012 Oct;27(10):1361-1367 [FREE Full text] [doi: 10.1007/s11606-012-2077-6] [Medline: 22618581]
62. Davison BJ, Breckon E. Factors influencing treatment decision making and information preferences of prostate cancer patients on active surveillance. Patient Educ Couns 2012 Jun;87(3):369-374. [doi: 10.1016/j.pec.2011.11.009] [Medline: 22177658]
63. Husson O, Mols F, van de Poll-Franse LV. The relation between information provision and health-related quality of life, anxiety and depression among cancer survivors: a systematic review. Ann Oncol 2011 Apr;22(4):761-772 [FREE Full text] [doi: 10.1093/annonc/mdq413] [Medline: 20870912]
64. Street Jr RL. Mediated consumer-provider communication in cancer care: the empowering potential of new technologies. Patient Educ Couns 2003 May;50(1):99-104. [doi: 10.1016/S0738-3991(03)00089-2] [Medline: 12767594]
65. Street Jr RL, Makoul G, Arora NK, Epstein RM. How does communication heal? Pathways linking clinician-patient communication to health outcomes. Patient Educ Couns 2009 Mar;74(3):295-301. [doi: 10.1016/j.pec.2008.11.015] [Medline: 19150199]
66. Kitchen PJ, Kerr G, Schultz DE, McColl R, Pals H. The elaboration likelihood model: review, critique and research agenda. Eur J Mark 2014 Nov 4;48(11/12):2033-2050. [doi: 10.1108/EJM-12-2011-0776]

67. Krebs P, Prochaska JO, Rossi JS. A meta-analysis of computer-tailored interventions for health behavior change. Prev Med 2010;51(3-4):214-221 [FREE Full text] [doi: 10.1016/j.ypmed.2010.06.004] [Medline: 20558196]
68. Lustria ML, Noar SM, Cortese J, van Stee SK, Glueckauf RL, Lee J. A meta-analysis of web-delivered tailored health behavior change interventions. J Health Commun 2013;18(9):1039-1069. [doi: 10.1080/10810730.2013.768727] [Medline: 23750972]
69. Noar SM, Benac CN, Harris MS. Does tailoring matter? Meta-analytic review of tailored print health behavior change interventions. Psychol Bull 2007 Jul;133(4):673-693. [doi: 10.1037/0033-2909.133.4.673] [Medline: 17592961]
70. Portnoy DB, Scott-Sheldon LA, Johnson BT, Carey MP. Computer-delivered interventions for health promotion and behavioral risk reduction: a meta-analysis of 75 randomized controlled trials, 1988-2007. Prev Med 2008 Jul;47(1):3-16 [FREE Full text] [doi: 10.1016/j.ypmed.2008.02.014] [Medline: 18403003]
71. Sohl SJ, Moyer A. Tailored interventions to promote mammography screening: a meta-analytic review. Prev Med 2007 Oct;45(4):252-261 [FREE Full text] [doi: 10.1016/j.ypmed.2007.06.009] [Medline: 17643481]
72. Ritterband LM, Thorndike FP, Cox DJ, Kovatchev BP, Gonder-Frederick LA. A behavior change model for internet interventions. Ann Behav Med 2009 Aug;38(1):18-27 [FREE Full text] [doi: 10.1007/s12160-009-9133-4] [Medline: 19802647]
73. Albada A, van Dulmen S, Spreeuwenberg P, Ausems MG. Follow-up effects of a tailored pre-counseling website with question prompt in breast cancer genetic counseling. Patient Educ Couns 2015 Jan;98(1):69-76. [doi: 10.1016/j.pec.2014.10.005] [Medline: 25455796]
74. Yun YH, Lee KS, Kim YW, Park SY, Lee ES, Noh DY, et al. Web-based tailored education program for disease-free cancer survivors with cancer-related fatigue: a randomized controlled trial. J Clin Oncol 2012 Apr 20;30(12):1296-1303. [doi: 10.1200/JCO.2011.37.2979] [Medline: 22412149]
75. Willems RA, Bolman CA, Mesters I, Kanera IM, Beaulen AA, Lechner L. Short-term effectiveness of a web-based tailored intervention for cancer survivors on quality of life, anxiety, depression, and fatigue: randomized controlled trial. Psychooncology 2017 Feb;26(2):222-230. [doi: 10.1002/pon.4113] [Medline: 26988800]
76. Jensen JD, King AJ, Carcioppolo N, Davis L. Why are tailored messages more effective? A multiple mediation analysis of a breast cancer screening intervention. J Commun 2012 Oct;62(5):851-868 [FREE Full text] [doi: 10.1111/j.1460-2466.2012.01668.x] [Medline: 26405350]
77. Kreuter MW, Oswald DL, Bull FC, Clark EM. Are tailored health education materials always more effective than non-tailored materials? Health Educ Res 2000 Jun;15(3):305-315. [doi: 10.1093/her/15.3.305] [Medline: 10977378]

Abbreviations

ELM: elaboration likelihood model
GIOCA: Gastro-Intestinal Oncological Centre Amsterdam
LC4MP: limited capacity model of motivated mediated message processing
NPIRQ: Netherlands Patient Information Recall Questionnaire
RCT: randomized controlled trial
STAI-6: 6-item version of the State-Trait Anxiety Inventory

Edited by G Eysenbach; submitted 16.04.19; peer-reviewed by J Noordman, T Dekkers, C Hoving; comments to author 07.06.19; revised version received 12.07.19; accepted 19.07.19; published 01.10.19.

Please cite as:
Nguyen MH, Smets EMA, Bol N, Loos EF, van Laarhoven HWM, Geijsen D, van Berge Henegouwen MI, Tytgat KMAJ, van Weert JCM
Tailored Web-Based Information for Younger and Older Patients with Cancer: Randomized Controlled Trial of a Preparatory Educational Intervention on Patient Outcomes
J Med Internet Res 2019;21(10):e14407
URL: https://www.jmir.org/2019/10/e14407
doi:10.2196/14407
PMID:31573911

Original Paper

Difference Between Users and Nonusers of a Patient Portal in Health Behaviors and Outcomes: Retrospective Cohort Study

Jing Huang[1*], PhD; Yong Chen[1,2,3,4,5*], PhD; J Richard Landis[1], PhD; Kevin B Mahoney[1,6], DBA

[1]Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, United States

[2]Institute for Biomedical Informatics, University of Pennsylvania, Philadelphia, PA, United States

[3]Center for Evidence-Based Practice, University of Pennsylvania, Philadelphia, PA, United States

[4]Center for Pharmacoepidemiology Research and Training, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, United States

[5]Applied Mathematics and Computational Science, University of Pennsylvania, Philadelphia, PA, United States

[6]Richard J Fox School of Business and Management, Temple University, Philadelphia, PA, United States

[*]these authors contributed equally

Corresponding Author:
Kevin B Mahoney, DBA
Perelman School of Medicine
University of Pennsylvania
3400 Civic Center Boulevard
Philadelphia, PA,
United States
Phone: 1 215 662 7393
Email: kevin.mahoney@uphs.upenn.edu

Abstract

Background: Patient portals are frequently used in modern health care systems as an engagement and communication tool. An increased focus on the potential value of these communication channels to improve health outcomes is warranted.

Objective: This paper aimed to quantify the impact of portal use on patients' preventive health behavior and chronic health outcomes.

Methods: We conducted a retrospective, observational cohort study of 10,000 patients aged 50 years or older who were treated at the University of Pennsylvania Health System (UPHS) from September 1, 2014, to October 31, 2016. The data were sourced from the UPHS electronic health records. We investigated the association between patient portal use and patients' preventive health behaviors or chronic health outcomes, controlling for confounders using a novel cardinality matching approach based on propensity scoring and a subsequent bootstrapping method to estimate the variance of association estimates.

Results: Patient-level characteristics differed substantially between portal users, comprising approximately 59.32% (5932/10000) of the cohort, and nonusers. On average, users were more likely to be younger (63.46 years for users vs 66.08 years for nonusers), white (72.77% [4317/5932] for users vs 52.58% [2139/4068] for nonusers), have commercial insurance (60.99% [3618/5932] for users vs 40.12% [1632/4068] for nonusers), and have higher annual incomes (US $74,172/year for users vs US $62,940/year for nonusers). Even after adjusting for these potential confounders, patient portal use had a positive and clinically meaningful impact on patients' preventive health behaviors but not on chronic health outcomes.

Conclusions: This paper contributes to the understanding of the impact of patient portal use on health outcomes and is the first study to identify a meaningful subgroup of patients' health behaviors that improved with portal use. These findings may encourage providers to promote portal use to improve patients' preventive health behaviors.

(J Med Internet Res 2019;21(10):e13146) doi:10.2196/13146

KEYWORDS

health behavior; health status; health care providers; patient portals

Introduction

Background

Chronic conditions are the leading causes of death and disability and key drivers of total health care costs in the United States [1-3]. According to the Centers for Disease Control and Prevention, 7 of the top 10 causes of death in the United States were chronic conditions. As of 2012, more than half of all the adults in the United States suffered from at least one chronic condition, with approximately 25% of adults having 2 or more chronic conditions; moreover, these proportions are expected to increase in the next decade [4-6]. Chronic conditions also account for the vast majority of health spending in the United States. Each year, 86% of the nation's total health care costs were for patients with at least one chronic condition and 71% were for patients with multiple conditions. In regard to public insurance, treatment of chronic conditions accounts for an even higher proportion of spending: 96% for Medicare and 83% for Medicaid [7-9].

From a public health perspective, primary prevention and screening for early-stage chronic conditions are considered the best strategies to prevent chronic conditions and to facilitate detection of disease at a milder stage of severity, thus incurring lower medical costs. Many chronic conditions could be prevented, delayed, or alleviated through simple lifestyle changes and other noninvasive interventions. In addition, it has been shown that preventive screening tests can reduce the death and comorbidity rates related to chronic conditions. For example, it has been reported that increased screening rates can reduce colorectal cancer deaths by 15% to 33% [10,11], and blood pressure checks for cardiovascular risk assessment can reduce cardiovascular morbidity of the population [12-14].

In the past few years, patient portals have been widely used in health care systems and have gained increasing attention for their potential values in improving health [15-22]. As a secure internet-based channel that provides patients with convenient access to personal health records, management of health services, and communication with health professionals, patient portals are considered as promising tools in promoting patient health outcomes, especially for chronic conditions, through promoting preventive behaviors, for example, taking screening tests, improving patient engagement in health outcomes, and facilitating self-management of chronic conditions [17,20,21].

Recently, several studies have investigated a variety of research questions related to portal use, including the characteristics of early portal users [23,24], the information being communicated through portals [25], patients' and clinicians' attitudes toward the use of patient portals [26], and the impact of patient portals on medication adherence and patient follow-up behaviors [20,27]. These important studies have contributed to the understanding of the role of patient portals in promoting communications between patients and health care providers in primary care and reducing health care cost by providing remote consultation as a low-cost alternative to physical office visits. Despite these successes, evidence is still scarce on whether the use of patient portals can ultimately improve patient health outcomes or modify patient preventive health behaviors positively [28,29].

Significance of This Study

A major challenge in evaluating the impact of the portal use on patient health outcomes using electronic health record (EHR) data is the inherent selection bias. Specifically, the use of patient portals can be related to patients' or health care providers' characteristics, which may confound the associations between the portal use and patient health outcomes. Naïve regression analysis of the health outcomes on patient portal use is biased upward if patients adopted the portal because they have higher health motivations and better strategies of managing health conditions.

As patient portals are expected to improve patient experience and engagement, we hypothesized that portal use may have a clinically meaningful impact on patients' preventive health behaviors, such as annual flu vaccination and blood pressure checks. Such an impact could be because of the benefit from better patient experience and engagement. Specifically, in this paper, we evaluated the impact of patient portal use on patients' preventive health behaviors and chronic health outcomes using EHR data from a well-defined patient cohort within the University of Pennsylvania Health System (UPHS).

Methods

Study Population

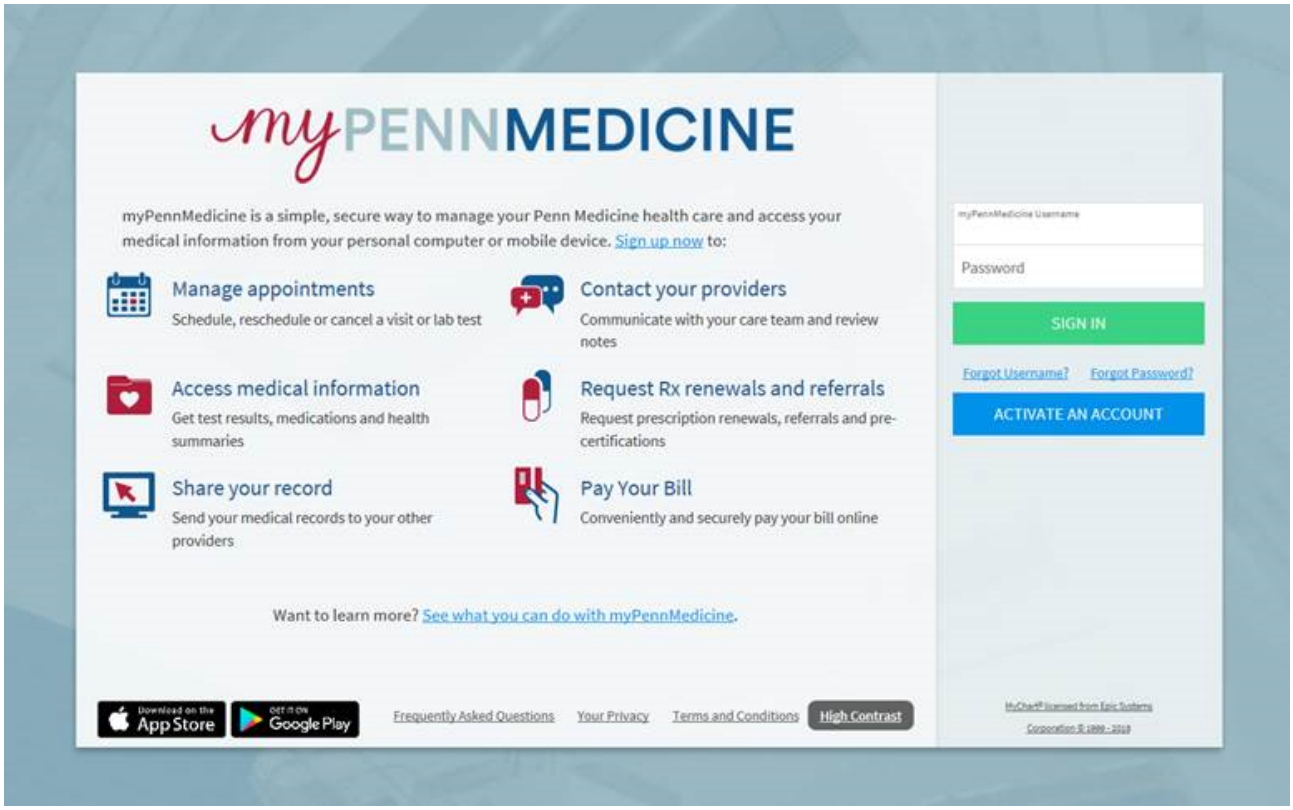
We conducted a retrospective observational cohort study using data from the Penn Data Store (PDS) [30]. There were 2 inclusion criteria for this study: (1) the patient was aged 50 years or older and (2) the patient had been seen by a Penn-employed primary care provider at least once within the time window from September 1, 2014, to October 31, 2016. The study population consisted of 10,000 patients randomly selected from the PDS. On the basis of these criteria, we assumed the study population was representative of the middle-aged patient population of greater Philadelphia. This study was reviewed and approved by the institutional review board at the University of Pennsylvania.

UPHS is a diverse research and clinical care organization located in Philadelphia, Pennsylvania. Founded in 1993, it currently operates under the direction and auspices of Penn Medicine, a division of the University of Pennsylvania. A total of 6 hospitals, including the flagship Hospital of the University of Pennsylvania, with over 5000 clinical care providers in the Greater Philadelphia Area, constitute the UPHS. UPHS serves over a million unique patients a year. Although most of these patients are located in the 28 counties in Southeast Pennsylvania, Central and Southern New Jersey, and Delaware, as a nationally recognized leader in care, UPHS provides care to patients in all 50 states, the District of Columbia, Puerto Rico, Guam, and the Virgin Islands. The average patient age is 50 years, and 60% of the patients are female. UPHS serves a diverse patient population where 62% of the patients are white, 21% are black or African American, 3% are Hispanic, and 3% are Asian. The patient payer mix is 44% commercial, 41% Medicare, 14% Medicaid, and 1% other.

myPennMedicine

myPennMedicine is a Penn Medicine–branded version of Epic's MyChart patient-facing electronic medical record. It is a patient portal that provides users with real-time information about medical records and test results, prescriptions, and appointments and other important health information. Patients may use the site to schedule appointments and laboratory tests, communicate with care teams, request prescription renewals and referrals, pay bills, and share records with other health care providers. It is available as a desktop Web portal and as an app for download from the Apple Store and Google Play. Patients must register, create an account, and log in to use these features. A screenshot of the *myPennMedicine* Web portal is shown in Figure 1. We defined patient portal users as patients who had registered for *myPennMedicine*. By May 2019, *myPennMedicine* had 591,784 unique and alive users. Among them, more than 66% had at least one activity in the past 365 days and approximately 45% and 28% had at least one activity in the past 90 and 30 days, respectively.

Figure 1. A screenshot of the myPennMedicine Web portal.



Health Outcomes of Interest

Health information of the study population was extracted from the PDS. The preventive health behaviors were measured by 4 binary (yes or no) variables contained in the EHR: annual flu vaccination, blood pressure check, colorectal cancer screen, and lipid level screen. We also calculated an overall summary variable, hereafter referred to as the composite prevention score, as the sum of the 4 binary variables, ranging from 0 to 4 (higher is better). For patients' chronic health outcomes, we studied 2 continuous and 2 binary variables from the EHR: systolic blood pressure, low-density lipoprotein, diabetes status (yes or no), and hypertension status (yes or no).

Patient- and Provider-Level Characteristics

We also extracted demographics of the patients and their health care providers from the PDS. Our analyses included 5 patient-level characteristics—age (in years), annual income (in US $), sex, race/ethnicity (white, black, Hispanic, and others), and insurance type (commercial, Medicaid, and Medicare)—and 3 provider-level characteristics—sex, research type (faculty vs nonfaculty), and certification type (physician vs advanced care provider).

Statistical Analysis

We evaluated the impact of the portal use on patient health outcomes by comparing the 4 preventive health behaviors and 4 chronic health outcomes between the portal users and nonusers. To control for potential confounders, including patient-level and health provider–level characteristics, we used a recently developed cardinality matching approach to match patient portal users to nonusers without replacement [31].

The cardinality matching procedure consists of 2 steps: balancing and pairing. First, this procedure maximizes the size of a match with prespecified requirements for balance on covariates, using integer programming. Specifically, we specified the differences in means of continuous covariates to be at most 0.05 SDs apart (moment balance) and required distributional balance on nominal covariates, without constraining users and nonusers to be matched within each category of each nominal covariate (fine balance) [32]. Then, with the groups determined and fixed, pairs were formed using

minimum distance pair matching for a robust rank-based Mahalanobis distance computed based on the propensity score [31]. The major advantage of the cardinality matching approach, compared with traditional matching approaches, is that it finds the maximal number of matched samples (ie, the maximum cardinality) with balanced covariates in the 2 groups as a whole. Traditional matching algorithms find matched groups that are balanced for covariates at the same time as they find pairs that are close in their covariates. By doing this, typical algorithms do not usually find the largest number of matched samples that balance observed covariates. More importantly, the cardinality matching method was shown to have greater efficiency and lower sensitivity to unmeasured biases [31-33].

In the second step of cardinality matching, that is, to find the matched pairs using propensity score–based distance, there are multiple options of distance measures. The commonly used distances calculated from the propensity score include Euclidean distance [34,35], weighted sum of absolute differences [36,37], and Mahalanobis distance [38-40]. In our analyses, we used a rank-based Mahalanobis distance with a caliper for penalty violations on the propensity score to constitute a robust distance for matching [31]. The rank-based Mahalanobis distance reduces the influence of outliers in the matching, and the penalty for caliper violations ensures good balance on propensity scores.

After matching, we evaluated the association between the portal use and patient health outcomes by comparing the difference in means for continuous variables and difference in proportions for binary variables, and we used a bootstrap method to compute the SEs of the estimates [41]. Although the standard testing procedures, for example, paired *t* test and McNemar test, have been widely used in practice to test the statistical significance of the association after propensity score matching, it has been shown that such procedures can be misleading with underestimated variances [42,43]. Specifically, in addition to sampling variation, the variance of the estimators after matching should also account for the variability because of the estimation of the propensity score, the imputation of the covariates, and possibly also the order in which individuals are matched [44]. In addition, standard nonparametric bootstrap procedures based on resampling fail to be consistent in this context [42] as the standard bootstrap procedure fails to reproduce the distribution of the true matching function. To properly account for the variation in the matching procedure, a large sample approximation to the normal distribution using the variance estimate suggested by Abadie and Imbens [45] or a novel bootstrap approach [41] can be used. We have adopted the latter strategy to obtain the variances of the effect size estimates after matching.

Results

Patient- and Provider-Level Characteristics

Among the 10,000 patients extracted from the PDS, a total of 5932 patients were registered to use the *myPennMedicine* patient portal. After matching, we obtained 3465 pairs (ie, 6930 patients) of patient portal users and nonusers. Table 1 summarizes the patient- and provider-level characteristics between the portal users and nonusers before the propensity score matching. We found that the patient-level characteristics were significantly different between the 2 groups. Compared with nonusers, users were more likely to be younger (63.46 years for users vs 66.08 years for nonusers; $P<.001$) and have higher income (US $74,172 for users vs US $62,940 for nonusers; $P<.001$). The percentage of white race in users was substantially higher among portal users (72.77% (4317/5932) for users vs 52.58% (2139/4068) for nonusers; $P<.001$). The percentage of payment by commercial insurance was also substantially higher (60.99% (3618/5932) for users vs 40.12% (1632/4068) for nonusers; $P<.001$), and the percentage of payment by Medicare or Medicaid was substantially lower (Medicare: 34.91% (2071/5932) for users vs 48.72% (1982/4068) for nonusers; $P<.001$ and Medicaid: 3.49% (207/5932) for users vs 10.08% (410/4068) for nonusers; $P<.001$). The difference in sex between users and nonusers was not statistically significant. We did not find a significant difference in any provider-level characteristic between the 2 groups. The statistical significance was adjusted using Bonferroni correction to control type I error.

We applied the cardinality matching procedure using variables of both patient- and provider-level characteristics listed in Table 1. After matching, we compared the summary statistics of these variables between the patient portal users and nonusers, shown in Table 2, and the variables were balanced.

Table 1. Summary of patient- and provider-level characteristics in patient portal users and nonusers before propensity score matching.

Variables	Users	Nonusers	*P* value
Patient characteristics[a]			
Age (years), mean (SD)	63.46 (9.22)	66.08 (11.37)	<.001
Income (US $), mean (SD)	74,171.99 (29,150.87)	62,939.74 (30,017.67)	<.001
Male, n (%)	2393 (40.34)	1592 (39.13)	.41
Race/ethnicity, n (%)			
White	4317 (72.77)	2139 (52.58)	<.001
Black	1134 (19.12)	1564 (38.45)	<.001
Hispanic	64 (1.08)	115 (2.83)	<.001
Other	417 (7.03)	250 (6.14)	.08
Insurance type, n (%)			
Commercial	3618 (60.99)	1632 (40.12)	<.001
Medicaid	207 (3.49)	410 (10.08)	<.001
Medicare	2071 (34.91)	1982 (48.72)	<.001
No insurance	36 (0.61)	44 (1.08)	<.001
Physician characteristics, n (%)			
Faculty	2035 (34.31)	1217 (29.92)	<.001
Physician	4466 (75.29)	3154 (77.53)	.01
Male	2458 (41.44)	1661 (40.83)	.53

[a]Categorical characteristics are given as the percentage of patients, and numerical characteristics are given as mean (SD).

Table 2. Summary of patient- and provider-level characteristics in patient portal users and nonusers after propensity score matching.

Variables	Users	Nonusers	*P* value
Patient characteristics[a]			
Age (years), mean (SD)	64.16 (9.97)	64.21 (10.00)	.89
Income (US $), mean (SD)	67,128 (29,609.13)	66,807 (30,726.53)	.62
Male, n (%)	1375 (39.68)	1380 (39.83)	.93
Race/ethnicity, n (%)			
White	2015 (58.15)	2002 (57.78)	.79
Black	1168 (33.71)	1184 (34.17)	.85
Hispanic	59 (1.70)	62 (1.79)	>.99
Other	223 (6.44)	217 (6.26)	.87
Insurance type, n (%)			
Commercial	1745 (50.36)	1719 (49.61)	.86
Medicaid	235 (6.78)	249 (7.19)	.74
Medicare	1464 (42.22)	1475 (42.57)	.72
No insurance	22 (0.64)	22 (0.63)	>.99
Physician characteristics, n (%)			
Faculty	1112 (32.09)	1116 (32.21)	>.99
Physician	2660 (76.77)	2661 (76.79)	.91
Male	1396 (40.29)	1390 (40.12)	>.99

[a]Categorical characteristics are given as the percentage of patients, and numerical characteristics are given as mean (SD).

Impact of Portal Use on Patient Health Outcomes

Table 3 presents the results from our analysis on matched patients. We found that patients' preventive health behaviors were significantly associated with portal use. Specifically, the proportions of annual flu vaccination, blood pressure check, and lipid level screen were substantially higher in portal users compared with nonusers (odds ratios, OR=1.58, 1.13, and 1.50, respectively; $P<.001$ for all 3 outcomes). The average composite prevention score was also significantly higher among portal users compared with nonusers (mean difference=0.22; $P<.001$). We also found that the proportion of colorectal cancer screening test between portal users and nonusers was statistically significant ($P<.001$), but the OR was very close to 1. The statistical significance could be because of the large sample size, but the estimated OR did not show any clinically meaningful difference. We did not find any clinically meaningful difference between patient portal users and nonusers in chronic health outcomes. In the Multimedia Appendix 1, we visualized the difference in the percentages of annual flu vaccination, blood pressure check, and lipid level screen and the mean of the composite prevention score between portal users and nonusers.

Table 3. Estimated difference of health outcomes between patient portal users and nonusers after propensity score matching.

Outcomes	n	Effect size	*P* value
Prevention health behaviors			
Flu shot	6930	1.58 (1.30 to 1.70)[a]	<.001
Blood pressure test	6930	1.13 (1.08 to 1.29)[a]	<.001
LDL[b] test	6930	1.50 (1.38 to 1.67)[a]	<.001
Colorectal cancer test	6930	0.99 (0.82 to 1.21)[a]	<.001
Composite prevention score (0-4)	6930	0.22 (0.18 to 0.26)[c]	<.001
Chronic health outcomes			
Systolic blood pressure	5692	−1.19 (−2.00 to 0.03)[c]	.06
LDL	3756	−2.34 (−3.65 to 1.20)[c]	.50
Diabetes status	6930	0.98 (0.79 to 1.20)[a]	.92
Hypertension status	6930	0.97 (0.84 to 1.12)[a]	.69

[a]Odds ratio (95% CI) for binary variables.

[b]LDL: low-density lipoprotein.

[c]Mean difference (95% CI) for continuous variables.

Discussion

Principal Findings and Comparison With Previous Work

We investigated the association between the use of patient portal and patient- or provider-level characteristics. We found large differences between the portal users and nonusers in patient-level characteristics but not in provider-level characteristics. As both patient- and provider-level characteristics can potentially confound the associations between portal use and patient health outcomes, we adjusted for these potential confounders by adopting a recently developed cardinality matching approach based on propensity scores. After matching, the patient- and provider-level characteristics between portal users and nonusers were found to be balanced. We then quantified the impact of the portal use on patient health outcomes by comparing 4 different preventive health behavior variables and 4 chronic health outcome variables between the 2 groups. To ensure valid CIs and *P* values for the effect size estimates, we adopted a novel bootstrap method to estimate correct variances of the estimated impact of the patient portal.

Using EHR data from UPHS, our study contributed independent evidence to the impact of patient portal use on health outcomes, especially on preventive health behavior outcomes. This new evidence can lead to a better understanding of the value of patient portals in promoting patient health care of the highest quality.

With the implementation of EHR systems in almost all large health care systems in the United States and the availability of patient portals in many health systems, quantifying the impact of the portal use on patient health outcomes is important to patients, health care providers, policy makers, and other stakeholders. Patient portals hold great promise in improving communication between health care providers and patients, improving patients' ability for self-management, enhancing their experience, improving health outcomes, and reducing medical costs. Recently, several investigators have published studies on understanding patient portal usability and satisfaction among users [15-29]. Investigators have found that the use of patient portals has significantly improved patients' ability to contact their providers directly, bypassing the usual gatekeepers in the practice, such as office staff and nurses in primary care; moreover, patient portals have also increased the follow-up rate in treatment of Crohn disease [17,20,25,27]. However, until now, very few studies have investigated the impact of portals on health outcomes. A very recent study of patient portals on

hospital outcomes found that patient portals might not change hospital outcomes, for example, 30-day readmissions, inpatient mortality, and 30-day mortality, by comparing hospitalized patients who used portals with those who did not [29].

As discussed in the study by Dumitrascu et al [29], an important area for future research was to investigate the impact of patient portal use on more immediate outcomes. Our study aimed to fill this important evidence gap by investigating the impact of portal use on patients' preventive health behaviors and chronic health outcomes using EHR data from UPHS. We chose preventive health behaviors and chronic health outcomes because they arguably create the greatest burden in health care and drive increases in medical costs. On the basis of our study on 10,000 adults (aged 50 years or older) recruited from an urban primary care system at Philadelphia, our investigation revealed that the use of patient portals significantly promoted preventive health behaviors, for example, taking flu vaccination and colorectal cancer screening tests, but it did not improve chronic health outcomes, for example, diabetes and hypertension status. We also found that the patients who have used patient portals have significantly different characteristics compared with patients who have never used patient portals. These patient portal users were more likely to be younger, have higher income, and use commercial insurance.

In this paper, we quantified the impact of patient portal use on preventive health behavior and chronic health outcomes based on the effect sizes (CIs) and their clinical meaningfulness. Given the large sample size of this study, we do not interpret statistical significance as clinical significance. For example, the mean difference of systolic blood pressure between portal users and nonusers was 1.19 mm Hg, which is close to statistically significant at .05 level but not clinically meaningful. In addition, interpretation of the study results may require caution in inferring causality. It should be noted that the diabetes and hypertension status in our study refer to the patients having the disease or not. The diseases could develop before or after the portal usage, such that for a subset of the study sample, the disease variables may be considered as baseline patient demographics rather than health outcomes. In this study, the comparisons of these disease statuses between portal users and nonusers are evaluations of cross-sectional associations between the portal usage and the disease status. Moreover, the observed preventive health behaviors may also be associated with specified chronic health outcomes. To reduce the potential confounder effect of the disease status on the association between the portal usage and prevention health behaviors, we also conducted the same analysis in subgroups defined by patients with hypertension, patients without hypertension, patients with diabetes, and patients without diabetes. We compared the same outcome variables between portal users and nonusers, except for the variable that was used to define the subgroup, and we obtained the same conclusion in these subgroup analyses as in the analysis using the entire dataset. We observed a statistically significant difference between portal users and nonusers in preventive health behaviors but not in chronic health outcomes.

Limitations

Our study also has a few limitations that deserve further investigation. First, the study focused on relatively healthy patients aged 50 years or older. The conclusion may not be applicable to younger patients or patients with severe disease. It would be interesting to see whether similar investigations conducted in different patient cohorts result in the same or different conclusions. Second, the time window for this study was constrained from 2014 to 2016. Future research with a longer time window could be conducted to study the temporal relationship between portal usage and the risk of developing chronic conditions. Moreover, the frequencies of portal usage among the portal users were unknown in this study. The estimated effects could be attenuated if the frequencies were very heterogenous across portal users. In this study, patient data with log-in statistics were not available to study the effect of portal usage frequency on health behaviors or outcomes.

Conclusions

The introduction of patient portals is widespread, and their success in promoting communication between patients and health care providers in primary care and reducing health care cost is being documented. This study is among the first to demonstrate that patient portal use is positively associated with patient preventive health behavior outcomes but not with chronic health outcomes. These findings contribute to the understanding and quantification of the impact of patient portal use on patient health outcomes. Additional research is required to confirm these findings. A future research direction is to understand the longitudinal impact of portal use on the trajectory of chronic health outcomes, which can provide new insights to patients, health care providers, policy makers, and other stakeholders on how patient portals can ultimately improve chronic health conditions.

Acknowledgments

This work was supported in part by the National Institutes of Health grants 1R01LM012607 (JH and YC), 1R01AI130460 (JH and YC), P50MH113840 (YC), 1R01AI116794 (YC), R01LM009012 (YC), and 1R01HD099348 (JH and YC). The authors wish to thank Elissa V Klinger, Master of Science, for editorial assistance.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Bar plot of percentage or mean values of prevention health behaviors among portal users and nonusers. BP: blood pressure; LDL: low-density lipoprotein.
[PDF File (Adobe PDF File), 70 KB - jmir_v21i10e13146_app1.pdf]

References

1. Bauer UE, Briss PA, Goodman RA, Bowman BA. Prevention of chronic disease in the 21st century: elimination of the leading preventable causes of premature death and disability in the USA. Lancet 2014 Jul 5;384(9937):45-52. [doi: 10.1016/S0140-6736(14)60648-6] [Medline: 24996589]
2. Mokdad AH, Marks JS, Stroup DF, Gerberding JL. Actual causes of death in the United States, 2000. J Am Med Assoc 2004 Mar 10;291(10):1238-1245. [doi: 10.1001/jama.291.10.1238] [Medline: 15010446]
3. Ward BW, Schiller JS, Goodman RA. Multiple chronic conditions among US adults: a 2012 update. Prev Chronic Dis 2014 Apr 17;11:E62 [FREE Full text] [doi: 10.5888/pcd11.130389] [Medline: 24742395]
4. Johnson NB, Hayes LD, Brown K, Hoo EC, Ethier KA, Centers for Disease Control and Prevention (CDC). CDC National Health Report: leading causes of morbidity and mortality and associated behavioral risk and protective factors--United States, 2005-2013. MMWR Suppl 2014 Oct 31;63(4):3-27. [Medline: 25356673]
5. Prince MJ, Wu F, Guo Y, Robledo LM, O'Donnell M, Sullivan R, et al. The burden of disease in older people and implications for health policy and practice. Lancet 2015 Feb 7;385(9967):549-562. [doi: 10.1016/S0140-6736(14)61347-7] [Medline: 25468153]
6. Milani RV, Lavie CJ. Health care 2020: reengineering health care delivery to combat chronic disease. Am J Med 2015 Apr;128(4):337-343. [doi: 10.1016/j.amjmed.2014.10.047] [Medline: 25460529]
7. Altman D, Frist WH. Medicare and medicaid at 50 years: perspectives of beneficiaries, health care professionals and institutions, and policy makers. J Am Med Assoc 2015 Jul 28;314(4):384-395. [doi: 10.1001/jama.2015.7811] [Medline: 26219056]
8. Barnett JC, Vornovitsky MS. United States Census Bureau. 2016. Health Insurance Coverage in the United States: 2015 - Current Population Reports URL: https://census.gov/content/dam/Census/library/publications/2016/demo/p60-257.pdf [accessed 2019-08-27]
9. DeNavas-Walt C, Proctor BD, Smith JC. United States Census Bureau. 2013. Income, Poverty, and Health Insurance Coverage in the United States: 2012 - Current Population Reports URL: https://www.census.gov/prod/2013pubs/p60-245.pdf [accessed 2019-08-27]
10. Zauber AG. The impact of screening on colorectal cancer mortality and incidence: has it really made a difference? Dig Dis Sci 2015 Mar;60(3):681-691 [FREE Full text] [doi: 10.1007/s10620-015-3600-5] [Medline: 25740556]
11. Siegel RL, Miller KD, Fedewa SA, Ahnen DJ, Meester RG, Barzi A, et al. Colorectal cancer statistics, 2017. CA Cancer J Clin 2017 May 6;67(3):177-193 [FREE Full text] [doi: 10.3322/caac.21395] [Medline: 28248415]
12. Mendis S. Cardiovascular risk assessment and management in developing countries. Vasc Health Risk Manag 2005;1(1):15-18 [FREE Full text] [doi: 10.2147/vhrm.1.1.15.58933] [Medline: 17319094]
13. Wallace ML, Ricco JA, Barrett B. Screening strategies for cardiovascular disease in asymptomatic adults. Prim Care 2014 Jun;41(2):371-397 [FREE Full text] [doi: 10.1016/j.pop.2014.02.010] [Medline: 24830613]
14. Patel P, Ordunez P, DiPette D, Escobar MC, Hassell T, Wyss F, Standardized Hypertension Treatment and Prevention Network. Improved blood pressure control to reduce cardiovascular disease morbidity and mortality: the standardized hypertension treatment and prevention project. J Clin Hypertens (Greenwich) 2016 Dec;18(12):1284-1294 [FREE Full text] [doi: 10.1111/jch.12861] [Medline: 27378199]
15. Weingart SN, Rind D, Tofias Z, Sands DZ. Who uses the patient internet portal? The PatientSite experience. J Am Med Inform Assoc 2006;13(1):91-95 [FREE Full text] [doi: 10.1197/jamia.M1833] [Medline: 16221943]
16. Sarkar U, Karter AJ, Liu JY, Adler NE, Nguyen R, López A, et al. Social disparities in internet patient portal use in diabetes: evidence that the digital divide extends beyond access. J Am Med Inform Assoc 2011 May 1;18(3):318-321 [FREE Full text] [doi: 10.1136/jamia.2010.006015] [Medline: 21262921]
17. Turvey C, Klein D, Fix G, Hogan TP, Woods S, Simon SR, et al. Blue Button use by patients to access and share health record information using the Department of Veterans Affairs' online patient portal. J Am Med Inform Assoc 2014;21(4):657-663 [FREE Full text] [doi: 10.1136/amiajnl-2014-002723] [Medline: 24740865]
18. Collmann J, Cooper T. Breaching the security of the Kaiser Permanente internet patient portal: the organizational foundations of information security. J Am Med Inform Assoc 2007;14(2):239-243 [FREE Full text] [doi: 10.1197/jamia.M2195] [Medline: 17213500]
19. Smith SG, O'Conor R, Aitken W, Curtis LM, Wolf MS, Goel MS. Disparities in registration and use of an online patient portal among older adults: findings from the LitCog cohort. J Am Med Inform Assoc 2015 Jul;22(4):888-895 [FREE Full text] [doi: 10.1093/jamia/ocv025] [Medline: 25914099]
20. Sarkar U, Lyles CR, Parker MM, Allen J, Nguyen R, Moffet HH, et al. Use of the refill function through an online patient portal is associated with improved adherence to statins in an integrated health system. Med Care 2014 Mar;52(3):194-201 [FREE Full text] [doi: 10.1097/MLR.0000000000000069] [Medline: 24374412]

21. Neuner J, Fedders M, Caravella M, Bradford L, Schapira M. Meaningful use and the patient portal: patient enrollment, use, and satisfaction with patient portals at a later-adopting center. Am J Med Qual 2015;30(2):105-113 [FREE Full text] [doi: 10.1177/1062860614523488] [Medline: 24563085]
22. Tavares J, Oliveira T. Electronic health record patient portal adoption by health care consumers: an acceptance model and survey. J Med Internet Res 2016 Mar 2;18(3):e49 [FREE Full text] [doi: 10.2196/jmir.5069] [Medline: 26935646]
23. Jung C, Padman R, Shevchik G, Paone S. Who are portal users vs early e-visit adopters? A preliminary analysis. AMIA Annu Symp Proc 2011;2011:1070-1079 [FREE Full text] [Medline: 22195168]
24. Mehrotra A, Paone S, Martich GD, Albert SM, Shevchik GJ. Characteristics of patients who seek care via e-visits instead of office visits. Telemed J E Health 2013 Jul;19(7):515-519 [FREE Full text] [doi: 10.1089/tmj.2012.0221] [Medline: 23682589]
25. Sun S, Zhou X, Denny JC, Rosenbloom TS, Xu H. Messaging to Your Doctors: Understanding Patient-provider Communications via a Portal System. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2013 Presented at: CHI'13; April 27-May 2, 2013; Paris, France p. 1739-1748. [doi: 10.1145/2470654.2466230]
26. Padman R, Shevchik G, Paone S, Dolezal C, Cervenak J. eVisit: a pilot study of a new kind of healthcare delivery. Stud Health Technol Inform 2010;160(Pt 1):262-266. [doi: 10.3233/978-1-60750-588-4-262] [Medline: 20841690]
27. Dykes D, Williams E, Margolis P, Ruschman J, Bick J, Saeed S, et al. Improving pediatric inflammatory bowel disease (IBD) follow-up. BMJ Qual Improv Rep 2016;5(1):u208961.w3675 [FREE Full text] [doi: 10.1136/bmjquality.u208961.w3675] [Medline: 27559472]
28. Goldzweig CL, Orshansky G, Paige NM, Towfigh AA, Haggstrom DA, Miake-Lye I, et al. Electronic patient portals: evidence on health outcomes, satisfaction, efficiency, and attitudes: a systematic review. Ann Intern Med 2013 Nov 19;159(10):677-687. [doi: 10.7326/0003-4819-159-10-201311190-00006] [Medline: 24247673]
29. Dumitrascu AG, Burton MC, Dawson NL, Thomas CS, Nordan LM, Greig HE, et al. Patient portal use and hospital outcomes. J Am Med Inform Assoc 2018 Apr 1;25(4):447-453. [doi: 10.1093/jamia/ocx149] [Medline: 29300961]
30. Perelman School of Medicine - University of Pennsylvania. Data Analytics Center: Penn Data Store URL: https://www.med.upenn.edu/dac/penn-data-store-warehouse.html [accessed 2019-08-27]
31. Zubizarreta JR, Paredes RD, Rosenbaum PR. Matching for balance, pairing for heterogeneity in an observational study of the effectiveness of for-profit and not-for-profit high schools in Chile. Ann Appl Stat 2014;8(1):204-231 [FREE Full text] [doi: 10.1214/13-AOAS713]
32. Rosenbaum PR. Design of Observational Studies. New York: Springer; 2010.
33. Kilcioglu C, Zubizarreta JR. Maximizing the information content of a balanced matched sample in a study of the economic performance of green buildings. Ann Appl Stat 2016 Dec;10(4):1997-2020. [doi: 10.1214/16-AOAS962]
34. Steiner PM, Cook D. Matching and propensity scores. In: Little TD, editor. The Oxford Handbook of Quantitative Methods. Volume 1: Foundations. Oxford: Oxford University Press; 2013:237-259.
35. Rosenbaum PR. Optimal matching for observational studies. J Am Stat Assoc 1989 Dec;84(408):1024-1032. [doi: 10.2307/2290079]
36. Guo S, Fraser MW. Propensity Score Analysis: Statistical Methods and Applications. First Edition. Thousand Oaks, CA: Sage Publications; 2010.
37. Austin PC. An introduction to propensity score methods for reducing the effects of confounding in observational studies. Multivariate Behav Res 2011 May;46(3):399-424 [FREE Full text] [doi: 10.1080/00273171.2011.568786] [Medline: 21818162]
38. Rosenbaum PR, Rubin DB. Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. Am Stat 1985 Feb;39(1):33-38. [doi: 10.2307/2683903]
39. D'Agostino Jr RB. Propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group. Stat Med 1998 Oct 15;17(19):2265-2281. [doi: 10.1002/(SICI)1097-0258(19981015)17:19<2265::AID] [Medline: 9802183]
40. Caliendo M, Kopeinig S. Some practical guidance for the implementation of propensity score matching. J Economic Surveys 2008 Feb;22(1):31-72. [doi: 10.1111/j.1467-6419.2007.00527.x]
41. Adusumilli K. Department of Economics - Iowa State University. 2017. Bootstrap Inference for Propensity Score Matching URL: https://www.econ.iastate.edu/files/events/files/karun_adusumilli_jmp.pdf [accessed 2019-08-27]
42. Abadie A, Imbens GW. On the failure of the bootstrap for matching estimators. Econometrica 2008;76(6):1537-1557. [doi: 10.3982/ECTA6474]
43. Abadie A, Imbens GW. Bias-corrected matching estimators for average treatment effects. J Business Econ Stat 2011 Jan;29(1):1-11. [doi: 10.1198/jbes.2009.07333]
44. Heckman JJ, Ichimura H, Todd P. Matching as an econometric evaluation estimator. Rev Econ Studies 1998 Apr 1;65(2):261-294. [doi: 10.1111/1467-937X.00044]
45. Abadie A, Imbens GW. Matching on the estimated propensity score. Econometrica 2016;84(2):781-807. [doi: 10.3982/ECTA11293]

Abbreviations

EHR: electronic health record
OR: odds ratio
PDS: Penn Data Store
UPHS: University of Pennsylvania Health System

Edited by G Eysenbach; submitted 14.12.18; peer-reviewed by X Wu, C Mamillapalli, K Wall, Y Yang, R Dendere; comments to author 23.03.19; revised version received 22.06.19; accepted 05.07.19; published 07.10.19.

Please cite as:
Huang J, Chen Y, Landis JR, Mahoney KB
Difference Between Users and Nonusers of a Patient Portal in Health Behaviors and Outcomes: Retrospective Cohort Study
J Med Internet Res 2019;21(10):e13146
URL: https://www.jmir.org/2019/10/e13146
doi:10.2196/13146
PMID:31593546

Original Paper

The Internet Intervention Patient Adherence Scale for Guided Internet-Delivered Behavioral Interventions: Development and Psychometric Evaluation

Fabian Lenhard[1], PhD; Kajsa Mitsell[2], MSc; Maral Jolstedt[1], MSc; Sarah Vigerland[1], PhD; Tove Wahlund[1], MSc; Martina Nord[1], MSc; Johan Bjureberg[1], PhD; Hanna Sahlin[1], PhD; Per Andrén[1], MSc; Kristina Aspvall[1], MSc; Karin Melin[3], RN, PhD; David Mataix-Cols[1], PhD; Eva Serlachius[1], PhD; Jens Högström[1], PhD

[1]Centre for Psychiatry Research, Department of Clinical Neuroscience, Karolinska Institutet, & Stockholm Health Care Services, Stockholm County Council, Stockholm, Sweden

[2]Stockholm Health Care Services, Stockholm County Council, Stockholm, Sweden

[3]Institute of Neuroscience and Physiology, The Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden

Corresponding Author:
Fabian Lenhard, PhD
Centre for Psychiatry Research
Department of Clinical Neuroscience
Karolinska Institutet, & Stockholm Health Care Services, Stockholm County Council
CAP Research Centre, Gävlegatan 22
SE-113 30
Stockholm,
Sweden
Phone: 46 8 514 522 12
Email: fabian.lenhard@ki.se

Abstract

Background: Patient adherence is defined as the extent to which a patient complies with medical or health advice. At present, there is a lack of reliable and valid measures specifically designed to measure adherence to internet-delivered behavioral interventions.

Objective: The objective of this study was to develop and psychometrically evaluate a novel measure of adherence to guided internet-delivered behavioral interventions.

Methods: In collaboration with experienced clinicians and researchers in the field, a 5-item, clinician-rated internet intervention Patient Adherence Scale (iiPAS) was developed. The initial scale was tested in a sample of children and adolescents (N=50) participating in internet-delivered cognitive behavioral therapy (ICBT) studies. A revised version of the iiPAS was then administered to a larger sample of children and adolescents (N=148) with various behavioral problems participating in ICBT trials. The scale was evaluated according to a classical test theory framework.

Results: The iiPAS demonstrated excellent internal consistency. Factor analyses revealed one underlying factor, explaining about 80% of the variance, suggesting that the scale captures a homogeneous adherence construct. The iiPAS was strongly associated with objective measures of patient activity in ICBT (number of logins, number of written characters, and completed modules). Furthermore, mid- and posttreatment ratings of the iiPAS were significantly correlated with treatment outcomes. By contrast, objective measures of patient activity in the Web-based platform did not correlate with treatment outcomes.

Conclusions: The iiPAS could be a useful tool to measure adherence in a broad range of internet-delivered behavioral interventions.

(J Med Internet Res 2019;21(10):e13602) doi:10.2196/13602

KEYWORDS

patient compliance; eHealth; measure; internet; cognitive behavioral therapy

Introduction

Background

Patient adherence, sometimes also termed compliance, is most broadly defined as "the extent to which the patient's behavior (in terms of taking medications, following diets, or executing other lifestyle changes) coincides with medical or health advice" [1]. Within the field of psychotherapy, adherence is usually conceptualized within the therapeutic tradition or according to the target condition that is being studied [2]. Accordingly, adherence is measured in various ways, for example, as qualitative or quantitative aspects of homework compliance [3], in-session engagement [4], or patient-therapist alliance [5]. Absence of patient adherence, on the other hand, is often referred to as resistance or ambivalence [6,7], and there are various therapeutic strategies that are specifically designed to target such resistance or ambivalence, such as motivational interviewing [8].

One important aspect of patient adherence is the logical assumption that adherent patients will have better treatment outcomes than nonadherent patients, though the evidence has not always supported this assumption. There is, for example, evidence that adherence to homework assignments in cognitive behavioral therapy (CBT) is associated with better outcomes, with a small to large effect of homework quality and quantity, but with substantial heterogeneity between studies [3,9]. Similarly, adherence to exposure and response prevention homework has been shown to be a strong predictor of short- and long-term outcome in CBT for obsessive-compulsive disorder (OCD) [10,11]. However, results from the field of pediatric anxiety have failed to establish a clear relationship between homework adherence and treatment outcome [12]. For example, in 1 study, neither the number of attended CBT sessions nor completed homework assignments predicted treatment outcome but a 1-item clinician-rated adherence measure did [13]. One study of CBT for anxiety disorders applied a combined measure of adherence, consisting of a weekly clinician rating of homework completion and a rating of engagement in the therapy session. The authors found that a combination of high treatment dose (high session attendance and completed exposure exercises) and high engagement predicted better treatment outcomes [14]. Importantly, it has been suggested that the adherence-outcome relationship could be affected by the quality of the adherence measure, and thus, reliable and valid measures are essential as research and clinical tools [15].

Internet-Delivered Interventions and Adherence

Internet-delivered behavioral interventions have emerged as an interdisciplinary development of psychotherapy and modern information technology. Internet-delivered behavioral interventions share common features as they mirror face-to-face psychotherapy regarding content and therapeutic techniques but are presented online via a personal login to a website or internet portal. The format usually resembles an online course, with instructive texts, pictures, videos, and written exercises, with the content structured similar to a book with chapters or modules. When therapist contact and feedback are involved, the treatment is referred to as guided or clinician-supported and otherwise as unguided or self-help [16]. Internet-delivered CBT (ICBT) has been one of the most prominent formats of this novel branch of interventions, with over 100 randomized controlled trials (RCTs) conducted in adults [17] and currently 19 RCTs in children and adolescents [18].

Adherence may be particularly difficult to measure in internet-delivered behavioral interventions [19]. In face-to-face treatment, many of the proposed active treatment components, such as behavioral or cognitive strategies, are conducted in-session together with a clinician. In internet-delivered behavioral interventions, the key therapy components are delivered online and practiced without the physical presence of the clinician. Depending on the intervention, the engagement in the active treatment components of the intervention may have low correlation with the number of completed chapters. For instance, patients may go on practicing their homework assignments without necessarily needing to log into the system periodically. Therefore, the number of completed chapters in internet-delivered behavioral interventions is not comparable with the number of completed sessions in face-to-face interventions and, thus, may not be a reliable measure of treatment adherence. As pointed out by Sieverink et al [20], the concept of "intended usage" could be a helpful extension of the traditional definition of adherence. Intended usage is defined as "the extent to which individuals *should* experience the content (of the intervention) to derive maximum benefit from the intervention, as defined or implied by its creators" [21].

Thus, defining and measuring the core therapeutic concepts and minimal effective use of those strategies may be a vital aspect of any measure of adherence to internet-delivered interventions. However, less than half of the studies in the field apply such a definition, and only few provide a justifiable definition of minimal required usage [20]. Currently, the most prevalent operationalizations of adherence are number of treatment modules completed by patients and number of logins [20,22]. The relationship between adherence and clinical outcome has thus far been limited by inconsistent reporting of adherence and a lack of agreement on measures [18,22].

In summary, there is an important research gap and a need for reliable and valid measures of patient adherence to internet-delivered interventions. We take the perspective that a useful measure of adherence to internet-delivered behavioral interventions should (1) not only capture the structural aspects of the internet format (such as completed chapters or modules) but also consider the intended usage of its central therapeutic aspects (such as use of behavioral strategies) and (2) include relevant, observable behaviors that represent the patients' adherence in a way that guides clinically meaningful decisions. Thus, the aim of this project was to construct and validate an adherence measure, evaluate its psychometric properties, and establish whether this new measure correlates better with treatment outcome than other traditional measures of adherence (eg, number of logins and completed chapters).

Methods

Scale Construction

A novel, clinician-rated measure called internet intervention Patient Adherence Scale (iiPAS) was constructed in a 4-step process following recommendations from the psychometric literature [23].

Step 1: Definition of Construct and Context

The first author (FL) conceptualized the scale and drafted a first version with the aim to measure patient adherence as one homogeneous construct. To ensure feasibility, usability, and precision, the scale was intended to be a short and time-efficient measure and assess patient behaviors that would be observable by clinicians. The context of the scale was defined as a clinician-rated measure to be used within clinician-supported, internet-delivered behavioral interventions. The advantage of a clinician-rated scale, compared with a self-rated measure, was justified by the ability of the clinician to assess the adherence of the patient relative to the intended, optimal use of the intervention, which was assumed crucial for the reliability of the instrument [20]. Building on the available evidence presented in the introduction, the scale should include items that capture structural aspects of internet-delivered interventions, equivalent to number of completed chapters, as well as content-related aspects, equivalent to the patients' engagement in the therapeutic content of the intervention .

Step 2: Response Format and Initial Item Formulation

To ensure optimal content validity, experienced clinicians and researchers with relevant competence (see "Expert Sample") were involved in choosing the response format and initial item formulation. The initial draft was circulated among the experts in 2 review rounds and edited iteratively by the first author according to suggestions and comments. The preliminary version of the iiPAS included 5 clinician-rated items rated on a 3-point Likert scale, including anchor labels for the end points of the scale (see description under "Measures").

Step 3: Preliminary Data Collection

The preliminary iiPAS was tested in a small sample ("Patient Sample 1") to provide basic psychometric information as well as to give clinicians the possibility to test the scale and give feedback regarding usability and applicability.

Step 4: Preliminary Examination of Psychometric Properties and Scale Revision

On the basis of additional feedback from step 3, the wording of the items was minimally revised for improved clarity. In addition, the 3-point Likert scale was extended to a 5-point scale to increase response variance. A preliminary psychometric analysis was carried out; results are presented below. For the full psychometric evaluation, the revised iiPAS was then administered to participants in several clinical trials of internet-delivered behavioral interventions ("Patient Sample 2").

Participants

Expert Sample

The expert sample consisted of 8 experienced researchers and clinicians who had been involved in clinical trials of internet-delivered behavioral interventions (coauthors FL, SV, MJ, MN, PA, KA, TW, and JH) within a Swedish research network called the Child Internet Project or Barninternetprojektet (BiP). The experts were involved in the initial scale development.

Patient Sample 1

Initial data were collected from 50 patients, of whom 33 (66%) were children participating in a study of ICBT for pediatric anxiety [24], and 17 (34%) were adolescents participating in a study of ICBT for pediatric OCD [25]. Main inclusion criteria were a primary diagnosis of separation anxiety disorder, social anxiety disorder, generalized anxiety disorder, panic disorder, or specific phobia according to *Diagnostic and Statistical Manual* of Mental Disorders, 4th edition (DSM-IV) and a clinical severity score above 3 on the Clinician Severity Rating (CSR) [26] or a primary diagnosis of OCD with a score of 16 or above on Children's Yale Brown Compulsive Scale [27]. Clinicians scoring the iiPAS were the same licensed clinical psychologists (N=13) who provided online support as a part of the ICBT interventions. For more detailed information, see the original publications of the respective trials [24,25]. The regional ethical review board in Stockholm, Sweden, approved both studies.

Patient Sample 2

Patient sample 2 consisted of 148 children and adolescents with various mental disorders, who together with their primary caretakers participated in different clinical trials. Common inclusion criteria within all trials were (1) age within the age interval for the respective trial (see "Interventions"), (2) internet access, (3) absence of suicidality or another severe condition that required immediate attention or higher priority, and (4) the clinical presentation of a mental disorder that was targeted in the specific trial (tic disorder, social anxiety disorder, generalized anxiety disorder, agoraphobia, panic disorder, specific phobia, separation anxiety disorder, OCD, and nonsuicidal self-injury). All conditions were diagnosed by experienced psychologists using semistructured diagnostic interviews based on DSM-5 criteria [28]. The iiPAS was scored by the treating clinicians (N=21), of whom 20 were licensed psychologists and a licensed nurse. For more detailed information, see the original publications of each trial [24,29-32]. The regional ethical review board in Stockholm, Sweden, approved all trials.

Interventions

All interventions were provided by researchers and clinicians associated to the BiP network, a collaboration between the Karolinska Institutet and the Child and Adolescent Mental Health Service in Stockholm, with a main research focus on development and evaluation of internet-delivered interventions for childhood mental disorders. All interventions in the current evaluation were clinician-guided ICBT interventions that also included parent support. Clinician advice, support, and feedback

were given predominantly via written messages within the secure internet platform or via occasional telephone calls. For a short summary of the interventions, see Table 1. More detailed descriptions can be found in the original publications [29-33].

Table 1. Summary of the included internet-delivered cognitive behavior therapy interventions.

Intervention	Target age (years)	Target condition	Number of modules/weeks	Main features
BiP[a] TIC [29]	7-17	Tourette syndrome/chronic tic disorder	10/10	Habit reversal training or exposure with response prevention
BiP SoFT[c] [30]	13-17	Social anxiety disorder	9/12	Exposure training and cognitive strategies
BiP OCD[b] [25,31]	7-17	OCD	12/12	Exposure with response prevention and parental coping strategies
BiP Anxiety [33]	8-12	Mixed anxiety disorders	12/12	Exposure training and child coping strategies
BiP ERITA[d][32]	13-17	Self-harm	11/12	Emotion regulation strategies and adaptive parental behaviors

[a]BiP: Barninternetprojektet.

[b]OCD: obsessive-compulsive disorder.

[c]SoFT: Social Fobi Tonåring.

[d]ERITA: Emotion Regulation Individual Therapy for Adolescent.

Procedure

The included clinical trials applied similar but not identical procedures. A description of the common aspects of the procedures is provided here; for a detailed description, see the respective trial publications. Participants in patient samples 1 and 2 were recruited via clinician- or self-referrals. Selection of eligible participants included both telephone screenings with a clinician and an in-person assessment with the child or adolescent and at least one caregiver. Fulfillment of diagnostic criteria was established by a clinician (most often a clinical psychologist) using the Mini-International Neuropsychiatric Interview child version (MINI Kid [34]) or the Anxiety Disorder Interviews Schedule Child and Parent versions (ADIS C/P [26]). Decision about inclusion or exclusion of participants was made after the in-person assessment, followed by treatment allocation.

The BiP TIC and BiP Anxiety trials had a randomized controlled design [24,29]. In BiP TIC, participants were randomized to either habit reversal training or exposure and response prevention. In BiP Anxiety, participants were randomized to either ICBT or an active control condition. The remaining trials had an open, within-group design.

The clinical outcome measures were assessed at pre- and posttreatment. Activity data from the use of the internet platform (number of logins, written characters, login time duration, and completed treatment chapters) were automatically collected during the treatments. The iiPAS was administered by the clinicians participating in the trials, on 2 separate occasions: halfway through the treatment (iiPAS-mid) and at posttreatment (iiPAS-post).

Measures

Internet Intervention Patient Adherence Scale

The iiPAS is a 5-item, clinician-rated measure of patient adherence to internet-delivered behavioral interventions. The 5 items are rated by treating clinicians and cover 5 central aspects of adherence: the patient's work pace, engagement, communication with the clinician, motivation for change, and login frequency. Each item is rated on a 0 to 4 Likert scale. Thus, the iiPAS-total score ranges from 0 to 20, with 0 indicating no adherence and 20 perfect adherence. The items are designed to generically suit different types of internet-delivered behavioral interventions. A prerequisite, however, is that the intervention is guided, that is, clinician-supported, as opposed to unguided or entirely self-help. The scale is constructed to be applicable to children and adolescents as well as adults. In the case where caregivers are involved in the patient's treatment, the iiPAS is primarily intended to evaluate the patient's adherence to treatment rather than the parent's or carer's involvement. See Multimedia Appendices 1 and 2 for a full copy of iiPAS in Swedish and English and scoring instructions. For permission requests to use or translate the scale, please contact the corresponding author.

Clinical Global Impression-Severity

Clinical Global Impression-Severity (CGI-S [35]) is a commonly used clinician-rated measure of global illness severity. Clinical severity is rated on a 7-point Likert scale, ranging from "normal, not ill" to "severely ill." The CGI-S is validated as a measure of clinical effectiveness within clinical trials as well as within routine clinical practice [36].

Clinician Severity Rating

CSR is rated in conjunction with the clinician-administered ADIS [26], which is a semistructured interview and is considered the gold standard in the assessment of pediatric anxiety disorders. The CSR is a clinician rating of clinical severity between 0 and 8, with 0 indicating absence of symptoms (or no disturbance in functioning and/or disability), 8 maximal severity, and 4 the cut-off for clinical severity (ie, meeting diagnostic criteria for the disorder). ADIS has previously shown excellent interrater reliability [37] and test-retest reliability [38].

Platform Usage Variables

The internet treatment platform that the patients logged in to for accessing treatment material (texts, videos, and exercises) provided user data on number of logins to the platform, duration of logged in time, and number of completed chapters and written characters by the patients in communication and exercises. The number of completed chapters was expressed in percentage relative to total number of chapters because of the different number of chapters in different interventions.

Statistical Analyses

Reliability Analyses

Cronbach alpha was calculated as a measure of internal consistency. An exploratory factor analysis (EFA) was conducted regarding the factor structure of the iiPAS, including the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy to determine the adequacy of the data for factor analysis as well as a Bartlett Test of Sphericity to determine the factorability of the five items.

Validity Analyses

Pearson correlations were used to explore convergent validity, defined as the association of the iiPAS with platform usage variables. In variables where outliers with large leverage were identified, Spearman correlation coefficient was used. The intercorrelations of iiPAS between the halfway treatment time point (iiPAS-mid) and at the end of treatment (iiPAS-post) were calculated. For some calculations, an iiPAS total score was used, which is the sum of iiPAS-mid and iiPAS-post, to obtain a measure of adherence that represents the whole treatment period from start to end of treatment. Regarding criterion validity, the association between iiPAS and symptom severity change from pre- to posttreatment was calculated. Because the available measures, CSR (in the BiP Anxiety group) and CGI-S (in the remaining groups), are similar clinician-rated, 1-item Likert scales, but not identical, the CSR and CGI-S change scores were z-transformed and then combined into a single new variable ("symptom severity change"), which was then further used in the analyses. There were only a few missing values in the dataset (2.7%), and missingness was, therefore, not further handled but listwise excluded from the respective analyses. Analyses were conducted in SPSS Statistics version 24 (IBM Corp) and R (R Core Group [39]).

Results

Study 1: Initial Evaluation

An initial version of the iiPAS was administered to 50 children and adolescents, aged between 7 and 17 years, receiving ICBT for OCD or anxiety disorders. Preliminary analyses of Cronbach alpha indicated excellent internal consistency (alpha=.93).

Study 2: Full Psychometric Evaluation

Sample Characteristics

The majority of the 148 participants were female (n=89, 60%), with a mean age of 12.7 years. Table 2 displays a detailed description of the characteristics of the participants in each trial.

Table 3 presents descriptive statistics for iiPAS at the 2 time points midtreatment and posttreatment and for the summary value, iiPAS-total, as well as platform usage variables.

Table 2. Sample characteristics (N=148).

Characteristic	BiP TIC (n=23)	BiP SoFT (n=30)	BiP OCD (n=24)	BiP Anxiety (n=46)	BiP ERITA (n=25)	Total
Age, years:months						
Mean (SD)	12:3 (2.56)	15:0 (1.22)	11:5 (2.75)	9:11 (1.36)	15:9 (1.31)	12:7 (2.94)
Min-max	8-17	13-17	8-17	8-12	13-17	8-17
Sex, n (%)						
Female	8 (35)	25 (83)	14 (58)	23 (50)	19 (76)	89 (60.1)
Male	15 (65)	5 (17)	10 (42)	23 (50)	1 (4)	55 (37.1)
Other/nonbinary sex	N/A[a]	N/A	N/A	N/A	5 (20)	4 (2.7)
Principal diagnosis, n (%)						
Tourette syndrome/chronic tic disorder	23 (100)	—[b]	—	—	—	23 (15.5)
Social anxiety disorder	—	30 (100)	—	5 (11)	—	35 (23.6)
OCD[c]	—	—	24 (100)	—	—	24 (16.2)
Separation anxiety disorder	—	—	—	17 (37)	—	17 (11.4)
Generalized anxiety disorder	—	—	—	12 (26)	—	12 (8.1)
Specific phobia	—	—	—	9 (20)	—	9 (6.0)
Panic disorder	—	—	—	3 (7)	—	3 (2.0)
Nonsuicidal self-injury	—	—	—	—	25 (100)	25 (16.8)

[a]N/A: not applicable.

[b]No occurrence.

[c]OCD: obsessive-compulsive disorder.

Table 3. Descriptive statistics for internet intervention Patient Adherence Scale and platform usage variables.

Variable	Mean (SD)	Total score percentiles				
		5th	25th	50th	75th	95th
iiPAS[a]-mid	13.43 (4.98)	5	10	15	17	20
iiPAS-post	11.27 (6.01)	0	7	12	16	19
iiPAS-total	24.74 (10.39)	7	17	26	34	39
Number of logins	32 (16.49)	—[b]	—	—	—	—
Logged in time (min)	2277[c] (5673.92)	—	—	—	—	—
Number of written characters	18,135[d] (50,624.17)	—	—	—	—	—
Completed chapters (%)	74.4 (25.8)	—	—	—	—	—

[a]iiPAS: internet intervention Patient Adherence Scale.

[b]Percentiles not calculated.

[c]Median=762.

[d]Median=5510.

Internal Consistency

Table 4 displays Cronbach alpha for iiPAS-mid and iiPAS-post across the included clinical studies. The internal consistency of the iiPAS was excellent at midtreatment as well as at posttreatment.

Table 4. Cronbach alpha for internet intervention Patient Adherence Scale.

Study	iiPAS[a]-mid	iiPAS-post
BiP[b] TIC	.90	.95
BiP SoFT[c]	.96	.96
BiP OCD[d]	.91	.92
BiP Anxiety	.96	.94
BiP ERITA[e]	.94	.97
All trials combined	.93	.95

[a]iiPAS: internet intervention Patient Adherence Scale.

[b]BiP: Barninternetprojektet

[c]SoFT: Social Fobi Tonåring

[d]OCD: Obsessive Compulsive Disorder

[e]ERITA: Emotion Regulation Individual Therapy for Adolescents

Factor Structure

For iiPAS-mid, a KMO=0.89 and Bartlett Test of Sphericity χ^2_{10}=578.5, *P*<.001 indicated factorability. The results of the EFA indicated a single factor with an eigenvalue of 3.97 that explained 79.3% of the variance. For iiPAS-post, KMO was 0.85 and Bartlett Test of Sphericity χ^2_{10}=787.0, *P*=.001, again indicating factorability. The EFA resulted in a single factor with an eigenvalue of 4.18 that explained 83.6% of the variance. Multimedia Appendix 3 presents communalities and factor loadings in both EFAs.

Convergent Validity

iiPAS-mid, iiPAS-post, and iiPAS-total scores were positively and significantly correlated with platform usage variables (*r*=0.25-0.79), apart from nonsignificant correlations between iiPAS and logged in time. There was also a substantial and significant correlation between iiPAS-mid and iiPAS-post (*r*=0.78). Visual inspection of the scatterplots did not indicate any nonlinear associations. Table 5 presents convergent validity correlations.

Criterion Validity

iiPAS-mid, iiPAS-post, and iiPAS-total scores correlated positively and significantly with the standardized symptom severity change score (*r*=0.17-0.19; Table 5). In contrast, number of logins, logged in time, written characters, and percentage of completed chapters were not significantly correlated with symptom severity change. A post hoc item-per-item analysis of the iiPAS-mid items regarding the symptom severity change score revealed a significant correlation with 2 of the 5 items, item 2 ("engagement with exercises"; *r*=0.21, *P*=.01), and item 4 ("motivation for change"; *r*=0.19, *P*=.02).

Table 5. Internet intervention Patient Adherence Scale correlations (*P* values) with platform usage variables and symptom severity change.

Measure	iiPAS[a]-mid (*P* value)	iiPAS-post (*P* value)	iiPAS-total (*P* value)	Number of logins (*P* value)	Logged in time (*P* value)	Written characters (*P* value)	Completed chapters (*P* value)
iiPAS-post	*.78[b] (<.001)*	—[c]	—	—	—	—	—
iiPAS-total	*.93 (<.001)*	*.95 (<.001)*	—	—	—	—	—
Number of logins	*.53 (<.001)*	*.57 (<.001)*	*.59 (<.001)*	—	—	—	—
Logged in time	.15 (.07)	.06 (.47)	.11 (.21)	*.19 (.03)*	—	—	—
Written characters	*.25 (.003)*	*.31 (<.001)*	*.31 (<.001)*	*.36 (<.001)*	.10 (.24)	—	—
Completed chapters	*.70 (<.001)*	*.79 (<.001)*	*.79 (<.001)*	*.58 (<.001)*	.09 (.29)	*.26 (<.001)*	—
Symptom severity change	*.17 (.04)*	*.19 (.02)*	*.18 (.04)*	.07 (.40)	−.04 (.61)	−.01 (.94)	.10 (.25)

[a]iiPAS: internet intervention Patient Adherence Scale.

[b]Italicized values indicating significant correlations at the <.05 level.

[c]Not applicable.

Discussion

Principal Findings

The objective of this project was to develop and psychometrically evaluate a novel measure of patient adherence in internet-delivered behavioral interventions. The measure was designed in close collaboration with experienced clinicians and researchers within the field and resulted in a clinician-rated, 5-item scale that was tested in a sample of children and adolescents with various behavioral disorders. Descriptive statistics indicated an appropriate range and variability of the iiPAS. The scale demonstrated excellent internal consistency. Consistently, factor analyses showed that the iiPAS measures one underlying factor, in agreement with the theoretical assumption that patient adherence within the context of internet interventions can be understood as one homogeneous construct.

Objective behavioral data were available through automatically registered internet platform data of number of logins, logged in time, written characters, and percentage of completed chapters. The iiPAS was significantly associated with those objective variables, apart from logged in time, which strongly suggests that the iiPAS indeed measures key aspects of patient adherence. The iiPAS was also significantly associated with symptom severity changes after treatment. Post hoc analyses revealed that 2 of the 5 items at midtreatment, engagement and motivation to change, were particularly associated with a decrease in symptoms. The midtreatment administration of the iiPAS may be especially interesting from a clinical point of view, as half of the internet treatment is still ahead. For example, a low iiPAS midtreatment score could inform the clinician that the adherence of the patient should be addressed for optimal outcome. The strong correlation of the iiPAS-mid and iiPAS-post suggests that little change in adherence is to be expected in the second half of the treatment. This could mean that other interventions such as motivational interventions might be needed to increase adherence, or that the intervention might need to be intensified or augmented with telephone, video calls or in-person therapist support.

Interestingly, none of the objective platform usage variables were associated with symptom severity change at posttreatment. This could possibly indicate that the mere adherence to the formal structure of an internet intervention (logging in, writing text, and completing chapters) is less important for the clinical outcome than actively engaging with the therapeutic content that is presented in the intervention (such as testing new behavioral or cognitive strategies and delivering homework exercises with high quality). The existing literature on patient adherence supports that quality of homework exercise completion is a better predictor of treatment outcome than number of completed exercises [10,13]. However, studies that are specifically designed to clarify the importance of adhering to the formal structure of a treatment in comparison with adhering to the therapeutic content are needed.

Limitations and Future Directions

Our analyses included diverse conditions, ranging from anxiety disorders, OCD, tics/Tourette's disorder, and self-harming behaviors. Future studies should aim for an even broader evaluation of the iiPAS in additional diagnoses and symptoms that are currently treated with internet interventions, especially in the behavioral medicine field, where about half of the internet treatment trials are conducted [18]. Moreover, although the scale was designed to be used across the life span, the age range of the participants is limited. Further validation work in adult samples is warranted. In addition, the iiPAS was evaluated in clinical trials in Sweden and, albeit, with projects from different research groups, all were located at the same university and used the same internet platform, limiting the generalizability to other contexts. Future studies should evaluate the iiPAS in other languages as well as in different cultural, health care, and academic environments.

Importantly, the iiPAS was designed to be useful in a wide range of internet-delivered, clinician-supported internet interventions of different therapeutic schools. However, the interventions in this study are all based on a CBT framework. An important next step would be to test the iiPAS in different internet-delivered behavioral interventions. Regarding the reliability analyses, an important aspect that was omitted in this study was the calculation of interrater reliability. Due to the scoring format of the iiPAS, this would require 2 clinicians for each patient, which was not feasible at the current stage of evaluation but would be important to take into account in future studies.

Finally, the sample size in this study was big enough for a global evaluation but too small to conduct detailed separate analyses for each specific condition or age group. A larger sample would allow analysis of the association between patient adherence and clinical outcomes in greater detail with regards to specific patient characteristics such as age, comorbidity, and symptom severity at baseline.

Conclusions

To summarize, the iiPAS demonstrated sound psychometric characteristics in a clinical sample of children and adolescents with a broad range of psychiatric and behavioral problems. The iiPAS appears to be a valuable tool to measure and study different aspects of adherence within the context of clinician-guided, internet-delivered behavioral interventions. Further evaluations within different age groups, types of internet-delivered interventions, and patient populations are warranted.

Acknowledgments

The project was funded by the Swedish Research Council for Health, Working Life and Welfare (Forte 2014-4052).

Authors' Contributions

FL initiated the conception of the iiPAS. FL, SV, MJ, MN, PA, KA, TW, JH, and DMC revised the iiPAS and contributed to the final version of the scale; FL, MJ, SV, TW, MN, JB, HS, PA, KA, and KM contributed to the data collection; FL, KM, and JH conducted the statistical analyses; and FL, KM, and JH drafted the initial manuscript. All authors contributed to the revisions of the manuscript and approved the final version.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Internet intervention Patient Adherence Scale—Swedish.
[PDF File (Adobe PDF File)69 KB - jmir_v21i10e13602_app1.pdf]

Multimedia Appendix 2
Internet intervention Patient Adherence Scale—English.
[PDF File (Adobe PDF File)161 KB - jmir_v21i10e13602_app2.pdf]

Multimedia Appendix 3
Supplemental tables—factor analyses: communalities and factor loadings.
[PDF File (Adobe PDF File)30 KB - jmir_v21i10e13602_app3.pdf]

References

1. Haynes BR, Taylor WD, Sackett DL. Compliance in Health Care. Baltimore, Maryland: Johns Hopkins University Press; 1979.
2. Nock MK, Ferriter C. Parent management of attendance and adherence in child and adolescent therapy: a conceptual and empirical review. Clin Child Fam Psychol Rev 2005 Jun;8(2):149-166. [doi: 10.1007/s10567-005-4753-0] [Medline: 15984084]
3. Mausbach BT, Moore R, Roesch S, Cardenas V, Patterson TL. The relationship between homework compliance and therapy outcomes: an updated meta-analysis. Cognit Ther Res 2010 Oct;34(5):429-438 [FREE Full text] [doi: 10.1007/s10608-010-9297-z] [Medline: 20930925]
4. Jungbluth NJ, Shirk SR. Therapist strategies for building involvement in cognitive-behavioral therapy for adolescent depression. J Consult Clin Psychol 2009 Dec;77(6):1179-1184. [doi: 10.1037/a0017325] [Medline: 19968393]
5. Martin DJ, Garske JP, Davis MK. Relation of the therapeutic alliance with outcome and other variables: a meta-analytic review. J Consult Clin Psychol 2000 Jun;68(3):438-450. [doi: 10.1037/0022-006X.68.3.438] [Medline: 10883561]
6. Beutler LE, Harwood TM, Michelson A, Song X, Holman J. Resistance/reactance level. J Clin Psychol 2011 Feb;67(2):133-142. [doi: 10.1002/jclp.20753] [Medline: 21108314]
7. Button ML, Westra HA, Hara KM, Aviram A. Disentangling the impact of resistance and ambivalence on therapy outcomes in cognitive behavioural therapy for generalized anxiety disorder. Cogn Behav Ther 2015;44(1):44-53. [doi: 10.1080/16506073.2014.959038] [Medline: 25277366]
8. Rubak S, Sandbaek A, Lauritzen T, Christensen B. Motivational interviewing: a systematic review and meta-analysis. Br J Gen Pract 2005 Apr;55(513):305-312 [FREE Full text] [Medline: 15826439]
9. Kazantzis N, Whittington C, Zelencich L, Kyrios M, Norton PJ, Hofmann SG. Quantity and quality of homework compliance: a meta-analysis of relations with outcome in cognitive behavior therapy. Behav Ther 2016 Sep;47(5):755-772. [doi: 10.1016/j.beth.2016.05.002] [Medline: 27816086]
10. Simpson HB, Maher MJ, Wang Y, Bao Y, Foa EB, Franklin M. Patient adherence predicts outcome from cognitive behavioral therapy in obsessive-compulsive disorder. J Consult Clin Psychol 2011 Apr;79(2):247-252 [FREE Full text] [doi: 10.1037/a0022659] [Medline: 21355639]
11. Simpson HB, Marcus SM, Zuckoff A, Franklin M, Foa EB. Patient adherence to cognitive-behavioral therapy predicts long-term outcome in obsessive-compulsive disorder. J Clin Psychiatry 2012 Sep;73(9):1265-1266 [FREE Full text] [doi: 10.4088/JCP.12l07879] [Medline: 23059155]
12. Hughes AA, Kendall PC. Prediction of cognitive behavior treatment outcome for children with anxiety disorders: therapeutic relationship and homework compliance. Behav Cogn Psychother 2007 May 25;35(4):487-494. [doi: 10.1017/s1352465807003761]
13. Lee P, Zehgeer A, Ginsburg GS, McCracken J, Keeton C, Kendall PC, et al. Child and adolescent adherence with cognitive behavioral therapy for anxiety: predictors and associations with outcomes. J Clin Child Adolesc Psychol 2019;48(Sup 1):S215-S226. [doi: 10.1080/15374416.2017.1310046] [Medline: 28448176]

14. Glenn D, Golinelli D, Rose RD, Roy-Byrne P, Stein MB, Sullivan G, et al. Who gets the most out of cognitive behavioral therapy for anxiety disorders? The role of treatment dose and patient engagement. J Consult Clin Psychol 2013 Aug;81(4):639-649 [FREE Full text] [doi: 10.1037/a0033403] [Medline: 23750465]
15. DiMatteo MR, Giordani PJ, Lepper HS, Croghan TW. Patient adherence and medical treatment outcomes: a meta-analysis. Med Care 2002 Sep;40(9):794-811. [doi: 10.1097/01.MLR.0000024612.61915.2D] [Medline: 12218770]
16. Andersson G. Using the internet to provide cognitive behaviour therapy. Behav Res Ther 2009 Mar;47(3):175-180. [doi: 10.1016/j.brat.2009.01.010] [Medline: 19230862]
17. Hedman E, Ljótsson B, Lindefors N. Cognitive behavior therapy via the internet: a systematic review of applications, clinical efficacy and cost-effectiveness. Expert Rev Pharmacoecon Outcomes Res 2012 Dec;12(6):745-764. [doi: 10.1586/erp.12.67] [Medline: 23252357]
18. Vigerland S, Lenhard F, Bonnert M, Lalouni M, Hedman E, Ahlen J, et al. Internet-delivered cognitive behavior therapy for children and adolescents: a systematic review and meta-analysis. Clin Psychol Rev 2016 Dec;50:1-10 [FREE Full text] [doi: 10.1016/j.cpr.2016.09.005] [Medline: 27668988]
19. Eysenbach G. The law of attrition. J Med Internet Res 2005 Mar 31;7(1):e11 [FREE Full text] [doi: 10.2196/jmir.7.1.e11] [Medline: 15829473]
20. Sieverink F, Kelders SM, van Gemert-Pijnen JE. Clarifying the concept of adherence to ehealth technology: systematic review on when usage becomes adherence. J Med Internet Res 2017 Dec 6;19(12):e402 [FREE Full text] [doi: 10.2196/jmir.8578] [Medline: 29212630]
21. Kelders SM, Kok RN, Ossebaard HC, van Gemert-Pijnen JE. Persuasive system design does matter: a systematic review of adherence to web-based interventions. J Med Internet Res 2012 Nov 14;14(6):e152 [FREE Full text] [doi: 10.2196/jmir.2104] [Medline: 23151820]
22. Donkin L, Christensen H, Naismith SL, Neal B, Hickie IB, Glozier N. A systematic review of the impact of adherence on the effectiveness of e-therapies. J Med Internet Res 2011 Aug 5;13(3):e52 [FREE Full text] [doi: 10.2196/jmir.1772] [Medline: 21821503]
23. Furr RM. Scale Construction and Psychometrics for Social and Personality Psychology. London, UK: Sage Publications; 2011.
24. Jolstedt M, Wahlund T, Lenhard F, Ljótsson B, Mataix-Cols D, Nord M, et al. Efficacy and cost-effectiveness of therapist-guided internet cognitive behavioural therapy for paediatric anxiety disorders: a single-centre, single-blind, randomised controlled trial. Lancet Child Adolesc Health 2018 Nov;2(11):792-801. [doi: 10.1016/S2352-4642(18)30275-X] [Medline: 30241993]
25. Lenhard F, Andersson E, Mataix-Cols D, Rück C, Vigerland S, Högström J, et al. Therapist-guided, internet-delivered cognitive-behavioral therapy for adolescents with obsessive-compulsive disorder: a randomized controlled trial. J Am Acad Child Adolesc Psychiatry 2017 Jan;56(1):10-9.e2 [FREE Full text] [doi: 10.1016/j.jaac.2016.09.515] [Medline: 27993223]
26. Silverman WK, Albano AM. Clinical Manual: Anxiety Disorders Interview Schedule for DSM-IV (Child and Parent Versions). New York: Oxford Press; 1996.
27. Scahill L, Riddle MA, McSwiggin-Hardin M, Ort SI, King RA, Goodman WK, et al. Children's Yale-brown obsessive compulsive scale: reliability and validity. J Am Acad Child Adolesc Psychiatry 1997 Jun;36(6):844-852. [doi: 10.1097/00004583-199706000-00023] [Medline: 9183141]
28. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders: DSM-5. Fifth Edition. Arlington, VA: American Psychiatric Publishing; 2013.
29. Andrén P, Aspvall K, de la Cruz LF, Wiktor P, Romano S, Andersson E, et al. Therapist-guided and parent-guided internet-delivered behaviour therapy for paediatric Tourette's disorder: a pilot randomised controlled trial with long-term follow-up. BMJ Open 2019 Feb 15;9(2):e024685 [FREE Full text] [doi: 10.1136/bmjopen-2018-024685] [Medline: 30772854]
30. Nordh M, Vigerland S, Öst LG, Ljótsson B, Mataix-Cols D, Serlachius E, et al. Therapist-guided internet-delivered cognitive-behavioural therapy supplemented with group exposure sessions for adolescents with social anxiety disorder: a feasibility trial. BMJ Open 2017 Dec 14;7(12):e018345 [FREE Full text] [doi: 10.1136/bmjopen-2017-018345] [Medline: 29247101]
31. Aspvall K, Andrén P, Lenhard F, Andersson E, Mataix-Cols D, Serlachius E. Internet-delivered cognitive behavioural therapy for young children with obsessive-compulsive disorder: development and initial evaluation of the BIP OCD junior programme. BJPsych Open 2018 May;4(3):106-112 [FREE Full text] [doi: 10.1192/bjo.2018.10] [Medline: 29971153]
32. Bjureberg J, Sahlin H, Hedman-Lagerlöf E, Gratz KL, Tull MT, Jokinen J, et al. Extending research on emotion regulation individual therapy for adolescents (ERITA) with nonsuicidal self-injury disorder: open pilot trial and mediation analysis of a novel online version. BMC Psychiatry 2018 Oct 11;18(1):326 [FREE Full text] [doi: 10.1186/s12888-018-1885-6] [Medline: 30305103]
33. Vigerland S, Ljótsson B, Thulin U, Öst LG, Andersson G, Serlachius E. Internet-delivered cognitive behavioural therapy for children with anxiety disorders: a randomised controlled trial. Behav Res Ther 2016 Jan;76:47-56 [FREE Full text] [doi: 10.1016/j.brat.2015.11.006] [Medline: 26649465]

34. Sheehan DV, Sheehan KH, Shytle RD, Janavs J, Bannon Y, Rogers JE, et al. Reliability and validity of the Mini International Neuropsychiatric Interview for Children and Adolescents (MINI-KID). J Clin Psychiatry 2010 Mar;71(3):313-326. [doi: 10.4088/JCP.09m05305whi] [Medline: 20331933]
35. Guy W. Assessment Manual for Psychopharmacology. Washington, DC: US Government Printing Office; 1976.
36. Berk M, Ng F, Dodd S, Callaly T, Campbell S, Bernardo M, et al. The validity of the CGI severity and improvement scales as measures of clinical effectiveness suitable for routine clinical use. J Eval Clin Pract 2008 Dec;14(6):979-983. [doi: 10.1111/j.1365-2753.2007.00921.x] [Medline: 18462279]
37. Ost LG, Svensson L, Hellström K, Lindwall R. One-session treatment of specific phobias in youths: a randomized clinical trial. J Consult Clin Psychol 2001 Oct;69(5):814-824. [doi: 10.1037/a0015158] [Medline: 11680558]
38. Silverman WK, Saavedra LM, Pina AA. Test-retest reliability of anxiety symptoms and diagnoses with the anxiety disorders interview schedule for DSM-IV: child and parent versions. J Am Acad Child Adolesc Psychiatry 2001 Aug;40(8):937-944. [doi: 10.1097/00004583-200108000-00016] [Medline: 11501694]
39. R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing; 2017.

Abbreviations

ADIS: Anxiety Disorders Interview Schedule
CBT: cognitive behavioral therapy
CGI-S: Clinical Global Impression-Severity
CSR: Clinician Severity Rating
DSM: Diagnostic and Statistical Manual of Mental Disorders
EFA: exploratory factor analysis
ICBT: internet-delivered cognitive behavioral therapy
iiPAS: internet intervention Patient Adherence Scale
KMO: Kaiser-Meyer-Olkin
OCD: obsessive-compulsive disorder
RCT: randomized controlled trial

Edited by G Eysenbach; submitted 08.02.19; peer-reviewed by E Gustafson, A Radomski, C Young; comments to author 28.03.19; revised version received 06.06.19; accepted 29.06.19; published 01.10.19.

Please cite as:
Lenhard F, Mitsell K, Jolstedt M, Vigerland S, Wahlund T, Nord M, Bjureberg J, Sahlin H, Andrén P, Aspvall K, Melin K, Mataix-Cols D, Serlachius E, Högström J
The Internet Intervention Patient Adherence Scale for Guided Internet-Delivered Behavioral Interventions: Development and Psychometric Evaluation
J Med Internet Res 2019;21(10):e13602
URL: https://www.jmir.org/2019/10/e13602
doi:10.2196/13602
PMID:31573901

Review

Effects of E-Learning in a Continuing Education Context on Nursing Care: Systematic Review of Systematic Qualitative, Quantitative, and Mixed-Studies Reviews

Geneviève Rouleau[1,2], RN, MSc; Marie-Pierre Gagnon[1,3], PhD; José Côté[2,4], RN, PhD; Julie Payne-Gagnon[1], RN, MA; Emilie Hudson[5], BSc, RN; Carl-Ardy Dubois[6,7], PhD; Julien Bouix-Picasso[4,8,9], CRNA, MPH

[1]Faculty of Nursing, Université Laval, Quebec, QC, Canada
[2]University of Montreal Hospital Research Centre, Montreal, QC, Canada
[3]Centre de Recherche sur les Soins et les Services de Première Ligne de l'Université Laval, Quebec, QC, Canada
[4]Faculty of Nursing, Université de Montréal, Montreal, QC, Canada
[5]School of Nursing, McGill University, Montreal, QC, Canada
[6]Public Health Research Institute, Université de Montréal, Montreal, QC, Canada
[7]Department of Management, Evaluation and Health Policy, School of Public Health, Université de Montréal, Montreal, QC, Canada
[8]Education and Health Practices Laboratory, Paris 13 University, Sorbonne Paris Cité University, Paris, France
[9]Department of Education for Non-Medical Personnel, French Military Health Service Academy, École du Val-de-Grâce, Paris, France

Corresponding Author:
Marie-Pierre Gagnon, PhD
Faculty of Nursing
Université Laval
Pavillon Ferdinand-Vandry 1050 Avenue de la Médecine
Quebec, QC, G1V 0A6
Canada
Phone: 1 418 656 2131 ext 407576
Email: marie-pierre.gagnon@fsi.ulaval.ca

Abstract

Background: E-learning is rapidly growing as an alternative way of delivering education in nursing. Two contexts regarding the use of e-learning in nursing are discussed in the literature: (1) education among nursing students and (2) nurses' continuing education within a life-long learning perspective. A systematic review of systematic reviews on e-learning for nursing and health professional students in an academic context has been published previously; however, no such review exists regarding e-learning for registered nurses in a continuing education context.

Objective: We aimed to systematically summarize the qualitative and quantitative evidence regarding the effects of e-learning on nursing care among nurses in a continuing education context.

Methods: We conducted a systematic review of systematic qualitative, quantitative, and mixed-studies reviews, searching within four bibliographic databases. The eligibility criteria were formulated using the population, interventions, comparisons, outcomes, and study design (PICOS) format. The included population was registered nurses. E-learning interventions were included and compared with face-to-face and any other e-learning interventions, as well as blended learning. The outcomes of interest were derived from two models: nursing-sensitive indicators from the Nursing Care Performance Framework (eg, teaching and collaboration) and the levels of evaluation from the Kirkpatrick model (ie, reaction, learning, behavior, and results).

Results: We identified a total of 12,906 records. We retrieved 222 full-text papers for detailed evaluation, from which 22 systematic reviews published between 2008 and 2018 met the eligibility criteria. The effects of e-learning on nursing care were grouped under Kirkpatrick's levels of evaluation: (1) nurse reactions to e-learning, (2) nurse learning, (3) behavior, and (4) results. Level 2, nurse learning, was divided into three subthemes: knowledge, skills, attitude and self-efficacy. Level 4, results, was divided into patient outcomes and costs. Most of the outcomes were reported in a positive way. For instance, nurses were satisfied with the use of e-learning and they improved their knowledge. The most common topics covered by the e-learning interventions were medication calculation, preparation, and administration.

Conclusions: The effects of e-learning are mainly reported in terms of nurse reactions, knowledge, and skills (ie, the first two levels of the Kirkpatrick model). The effectiveness of e-learning interventions for nurses in a continuing education context remains unknown regarding how the learning can be transferred to change practice and affect patient outcomes. Further scientific, methodological, theoretical, and practice-based breakthroughs are needed in the fast-growing field of e-learning in nursing education, especially in a life-learning perspective.

Trial Registration: International Prospective Register of Systematic Reviews (PROSPERO) CRD42016050714; https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=50714

(J Med Internet Res 2019;21(10):e15118) doi:10.2196/15118

KEYWORDS

continuing education; e-learning; nurses; nursing care; systematic review of systematic reviews

Introduction

Background

E-learning is rapidly growing as an alternative way of delivering education [1,2]. Nicoll et al [3] used the term *technology-enhanced learning* and stated that "it is a means by which learners can be provided with enhanced or transformed educational experiences." Many other terms have been used synonymously and interchangeably to designate e-learning, such as computer-assisted learning, online learning, or Web-based learning [4]. For the purpose of this paper, we will use *e-learning* as an umbrella term to entail a variety of electronic, digital, or mobile devices used to support learning [5]. Clark and Mayer [6] specify elements about the *what*, *how*, and *why* of e-learning. The *what* includes content and instructional methods. The *how* encompasses elements such as the format (eg, asynchronous and webinars) and the use of multimedia (eg, video, animation, and printed words). The *why* is about, for instance, the achievement of learning objectives and/or the performance of skills applied in a workplace context.

In the literature targeting the use of e-learning in nursing, two populations and contexts are discussed. The first one is education among nursing students (eg, in Voutilainen et al [7]) who participate in educational programs mainly offered in academic settings. For instance, undergraduate nursing students have to develop entry-level competencies to meet the practice expectations required in getting their registered nurse (RN) licensure in order to "provide safe, competent, compassionate, and ethical nursing care in a variety of practice settings" [8]. The second context is continuing education (CE), also called continuing professional development [9] or continuing competency [8], targeting a life-long learning perspective and staff development (eg, Knapp and Byers [10]). RNs have to meet CE expectations to be eligible to renew their licensure and registration each year, with the goals of acquiring new competencies, maintaining the acquired ones, enhancing their practice, and keeping their skills relevant and up-to-date [8]. We refer here to CE programs that are applicable in workplace settings. The use of e-learning by nurses in a CE context is the one that retained our attention [11] for two main reasons: much more attention has been given to nursing students than to RNs [5,12] and this CE context will be informative to lay the groundwork for a wider research project.

A previous systematic review of systematic reviews (SRSRs) of e-learning for nursing and health professional students in an academic context has been conducted [5,12]. The findings show that e-learning is equivalent to traditional learning. However, e-learning has proven to have large effects compared to no intervention in health professions [13]. To the best of our knowledge, we found no SRSRs about e-learning used by RNs in a CE context.

Objective

We aimed to systematically summarize the qualitative and quantitative evidence that comes from systematic qualitative, quantitative, and mixed-studies reviews regarding the effects of e-learning on nursing care among nurses in a CE context.

Methods

The protocol was registered in the International Prospective Register of Systematic Reviews (PROSPERO) (number: CRD42016050714) and was published elsewhere [11].

Design

We conducted a systematic review (SR) of systematic qualitative, quantitative, and mixed-studies reviews with the intent of bringing together, summarizing, and enhancing accessibility of existing evidence [14]. We combined outcomes from various types of SRs and synthesized qualitative and quantitative evidence. This type of synthesis is useful in identifying existing e-learning interventions used by RNs in their workplace settings and in describing the range of outcomes of interest measured, documented, and informed by the Nursing Care Performance Framework (NCPF).

We used the Cochrane Collaboration methodology [15] and other relevant works in this domain [16,17] to structure and report the SRSRs.

Nursing Care Performance Framework

We planned to use the NCPF to guide our synthesis. The NCPF is based on Henriksen et al's work [18], which depicts a conception of nursing care as a complex, whole entity; this entity is encompassed by many interrelated and interdependent subsystems and components that are logically coordinated and oriented toward the achievement of optimal outcomes for patients [19]. The NCPF is a systemic and organizational model aimed to measure the performance of nurses in the health care system through a set of indicators sensitive to various aspects

of nursing care. The rationale for using the NCPF was based on our previous work [20]. We conducted an SRSRs of the effects of information and communications technologies on nursing care. We then categorized these effects based on the following nursing care subsystems pertaining to the NCPF: nursing care structure (eg, nursing staff supply and profiles, work conditions, and nursing staff stability), nursing services (eg, professional practice environment, nursing processes, and interventions), and patient outcomes (eg, patient functional status and care safety). Our first intent in this current SRSRs was to use the NCPF for guidance and as a starting point for data extraction and analysis, while remaining open to the emergence of new data (ie, outcomes) that were not part of the framework. We expected to get a comprehensive portrait of how dimensions and indicators of nursing care, as developed in the NCPF, could be influenced by the use of e-learning interventions in a nursing CE context. In other words, we identified, from the NCPF, a pre-established range of possible outcomes and indicators related to nursing care, for which data would be sought in this SRSRs [11].

Eligibility Criteria

The scope was formulated using the population, intervention, comparison, outcomes, and study design (PICOS) format [15,21]. Eligibility criteria are presented in Table 1.

Table 1. Eligibility criteria.

SR[a] components	Inclusion criteria	Exclusion criteria
Population	RNs[b], according to the professional legislation of each country	Undergraduate nursing students in an academic context
Intervention	E-learning (ie, use of electronic, digital, or mobile devices to support learning) used in a continuing education context	Any type of simulation with a physical mannequin
Comparison	Face-to-face learning, any other e-learning intervention, or blended learning	N/A[c]
Outcomes	Primary outcomes: effects of e-learning on nursing care, including (1) nursing resources (eg, working conditions, nursing staff supply, and staff maintenance) and (2) nursing services (eg, nurses' practice environments, nursing processes and interventions, and professional satisfaction) Secondary outcomes: Effects of e-learning on patient outcomes (eg, patients' empowerment, comfort, and quality of life)	N/A
Study design	Systematic qualitative, quantitative, and mixed-studies reviews published in French, English, or Spanish	Grey literature and non-SR, such as literature reviews

[a]SR: systematic review.

[b]RN: registered nurse.

[c]N/A: not applicable.

Search Strategy and Selection Criteria

We searched for articles that were published between January 1, 2006, and January 26, 2017, in the following electronic databases: PubMed, Embase, Cumulative Index of Nursing and Allied Health Literature (CINAHL), and Joanna Briggs Institute (JBI). We updated the search strategy to include articles published between January 1, 2017, and November 11, 2018. The search strategy time frame was partially informed by the work of de Caro et al [5,12], who conducted an SRSRs on a similar topic. They performed search strategies on articles published between 2003 and 2013. Included SRs were published between 2008 and 2013. We extended our search strategy to include articles published from 2006 onward to capture SRs that could have been missed in previous work.

We developed a structured search strategy that was validated by a health information specialist. We used the thesaurus terms from each database and used free text to target the *title* and *abstract* fields. An example of the search strategy in PubMed has been presented elsewhere [11]. For the JBI database, we hand searched the whole database for relevant literature. We collected the results of each database search in EndNote reference manager, version X7.7.1 (Clarivate Analytics) and we removed duplicate citations. Furthermore, we hand searched for relevant SRs, contacted authors to find other relevant works in this domain, and consulted the reference lists of included SRs.

Selection of Systematic Reviews

We used DistillerSR (Evidence Partners), a Web-based SR software, to perform the screening and selection of SRs as well as the data extraction. A group of three reviewers (GR, JPG, and EH) independently screened the title and abstract of the papers in order to assess their eligibility. Each paper was reviewed twice, by two of the three reviewers. When consensus was not reached, arbitration was done with the third review author who was not involved in the selection of a specific SR. After the first round of screening, we retrieved full-text copies of publications that met the pre-established inclusion criteria and we assessed them twice.

Methodological Quality Assessment of Included Systematic Reviews

The methodological quality assessment is defined as the critical appraisal of each SR and the extent to which authors of each SR met the highest possible standards in conducting and reporting their research process. Methodological quality also refers to risk of bias (ie, deviations of findings from the truth); flaws in design, conduct, analysis, and/or reporting can be the cause of these deviations [17].

Methodological quality was done independently by two reviewers (GR and JBP) using two critical appraisal tools in order to cover a wider and complementary range of criteria: Assessment of Multiple Systematic Reviews (AMSTAR) 2 [22] and Risk Of Bias In Systematic Reviews (ROBIS) [23]. AMSTAR 2 is a 16-item instrument that provides detailed and comprehensive assessment of SRs that include randomized or nonrandomized studies of health care interventions. ROBIS contains 21 signaling questions divided into three phases:

1. The assessment of relevance (optional).
2. The identification of concerns with the review process, in which bias can be introduced from within four domains:
   a. Study eligibility criteria.
   b. Identification and selection of studies.
   c. Data collection and study appraisal.
   d. Synthesis and findings.
3. Overall judgment about risk.

These tools are best suited for quantitative SRs and were not designed for systematic qualitative and mixed-studies reviews. However, because there was no consensus on how to assess methodological quality of qualitative and mixed-studies reviews at the time we began the SRSRs, we used both tools: AMSTAR 2 and ROBIS. Any disagreements that arose between the reviewers during the methodological quality assessment process were resolved through discussion.

Data Extraction

A team of three authors (GR, JPG, and EH) conducted data extraction. Each paper was extracted independently by two of the three reviewers. We extracted the following data: (1) general characteristics of the SRs (eg, purpose, type of SR, number of primary studies included, populations, and settings); (2) details about e-learning interventions, comparisons, and the use of theories (ie, in the development and evaluation processes); and (3) outcomes, including their nature (ie, qualitative and/or quantitative) and direction (ie, positive, negative, or no effect). We used the adapted version of the NCPF [19] as a guide to extract outcomes. The dual extraction of outcomes data is particularly important, since these data are directly used in synthesizing the evidence that informs the conclusions of the review [23].

As recommended by Higgins et al [24], all the steps mentioned before (ie, selection of SRs, methodological quality assessment, and data extraction) were done by two authors working independently in order to minimize the risk of making mistakes and of being influenced by a single person's biases. In total, four authors were involved in performing these steps (GR, JPG, EH, and JBP).

Data Synthesis

The first author (GR) performed a qualitative thematic synthesis using a data-based convergent synthesis design [25,26]. We qualified quantitative data by using a narrative synthesis to describe the effect of e-learning on nursing care. We transformed the numerical data in specific themes and subthemes. This approach in conducting the data synthesis was chosen considering the mixed nature of evidence (ie, qualitative and quantitative) and the exploratory lens of this SRSRs. Two reviewers (EH and JPG) were involved in validating the data synthesis.

Overlap in Systematic Reviews

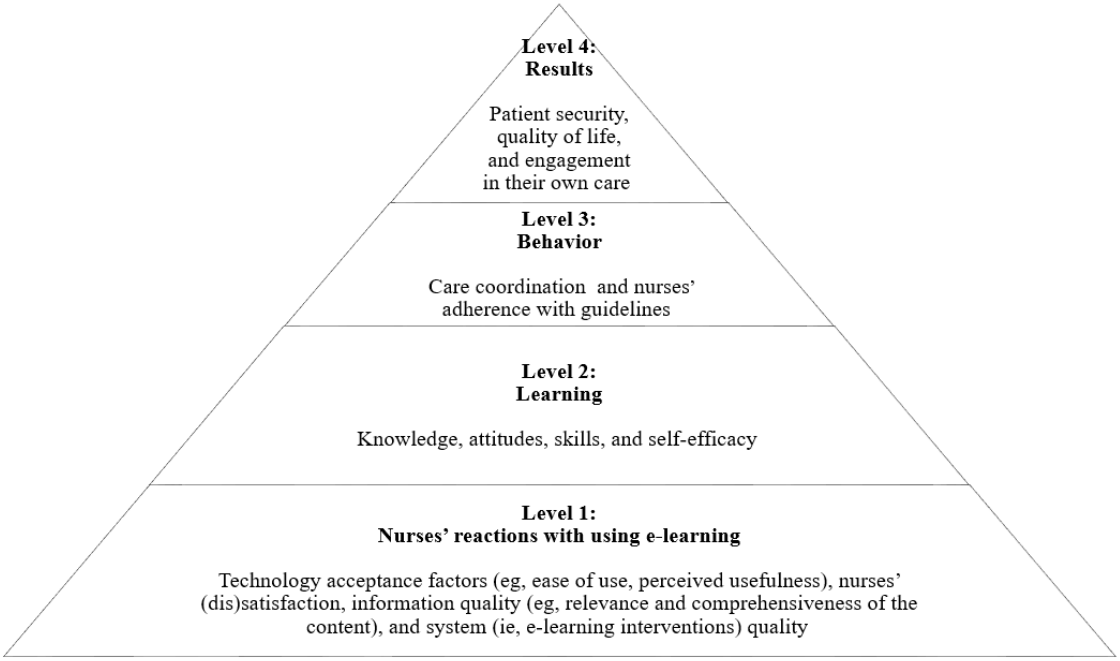
As mentioned in the protocol [11], one of the challenges encountered when conducting SRSRs is the identification of overlap in SRs [27-29] (ie, "when the primary studies included within the SRs had multiple related publications that were referenced differently across SRs" [29]). Authors of SRSRs need to closely examine the content of the SRs and their included primary studies to accurately assess the extent and nature of the overlap. The ways of managing overlap in SRs depend on the purpose of the SRSRs and the method of data analysis [17]. Pollock et al [17] suggest that when the purpose of the SRSRs is to present and describe the body of knowledge that comes from SRs, it may be appropriate to include the results of all relevant SRs, regardless of the overlap across primary studies. Considering the exploratory lens of our SRSRs and our intent to draw a broad picture of the effects of e-learning interventions used by nurses in a CE context, we created a citation matrix [17] (see Multimedia Appendix 1) to visually represent the overlap between primary studies within included SRs. The implications of these overlaps do not have consequences on, for example, overestimating the effects of e-learning interventions that would bias recommendations to use a specific intervention over the others. The purpose of this SRSRs is not prescriptive and is not intended to inform or guide decision making, policy, or practice recommendations.

Deviation From the Protocol

As previously mentioned, we used the NCPF to extract and classify the outcomes but we did not use it, as we had planned, to synthesize data regarding the effects of e-learning on nursing care. Indeed, this framework offers a broad perspective of the nursing care system, considering the diversity of nursing-sensitive indicators that are centered on structure and resources, nursing services and processes, and patient outcomes. However, most of these indicators do not reflect the current state of knowledge deriving from the effects of e-learning reported in the literature. These effects are rather circumscribed around nurses' level of satisfaction, knowledge, or skills acquisition, which fit more with the Kirkpatrick model [30]. This model proposes four distinct levels as a sequence of ways to evaluate the effectiveness of an educational program: (1) reactions, (2) learning, (3) behavior, and (4) results. Level 1 (ie, reactions) is about nurses' satisfaction with e-learning interventions. Level 2 (ie, learning) refers to the extent to which

nurses change or improve attitudes, knowledge, skills, and/or self-efficacy as a result of attending the e-learning interventions [30]. Level 3 (ie, behavior) is the extent to which nurses' learning has been translated into their postlearning behavior or their clinical performance [9]. Level 4 (ie, results) can be seen as patients' health outcomes resulting from the influence of e-learning interventions on nurses' behavior changes, which was adapted from Légaré et al [9]; this level can also be seen as other outcomes, such as costs [30]. In Figure 1, the Kirkpatrick model is presented, supported by some concrete examples provided by Shen et al [31].

Figure 1. The Kirkpatrick model.



Other frameworks could have been selected to extract and synthesize data, including the Expanded Outcomes Framework [32] or the Jeffries simulation model [33], since the *outcomes* component of the latter model can be potentially applicable to e-learning. However, we chose the Kirkpatrick model because it is well documented and extensively used in many educational contexts, including e-learning in the CE context [31,34].

Once we performed the first extraction using the NCPF, the data were read several times by three authors (GR, EH, and JPG). The first author built a thematic tree based on the reading of all material through line by line coding (ie, the inductive part) and based on existing works [19,30] (ie, the deductive part). This SRSRs was then guided by these two models [19,30] at different points in time: the use of the NCPF [19] was preplanned, and the use of the Kirkpatrick model [30] was decided during the process of data analysis and synthesis. The presentation of findings are supported by the four levels of evaluation [30].

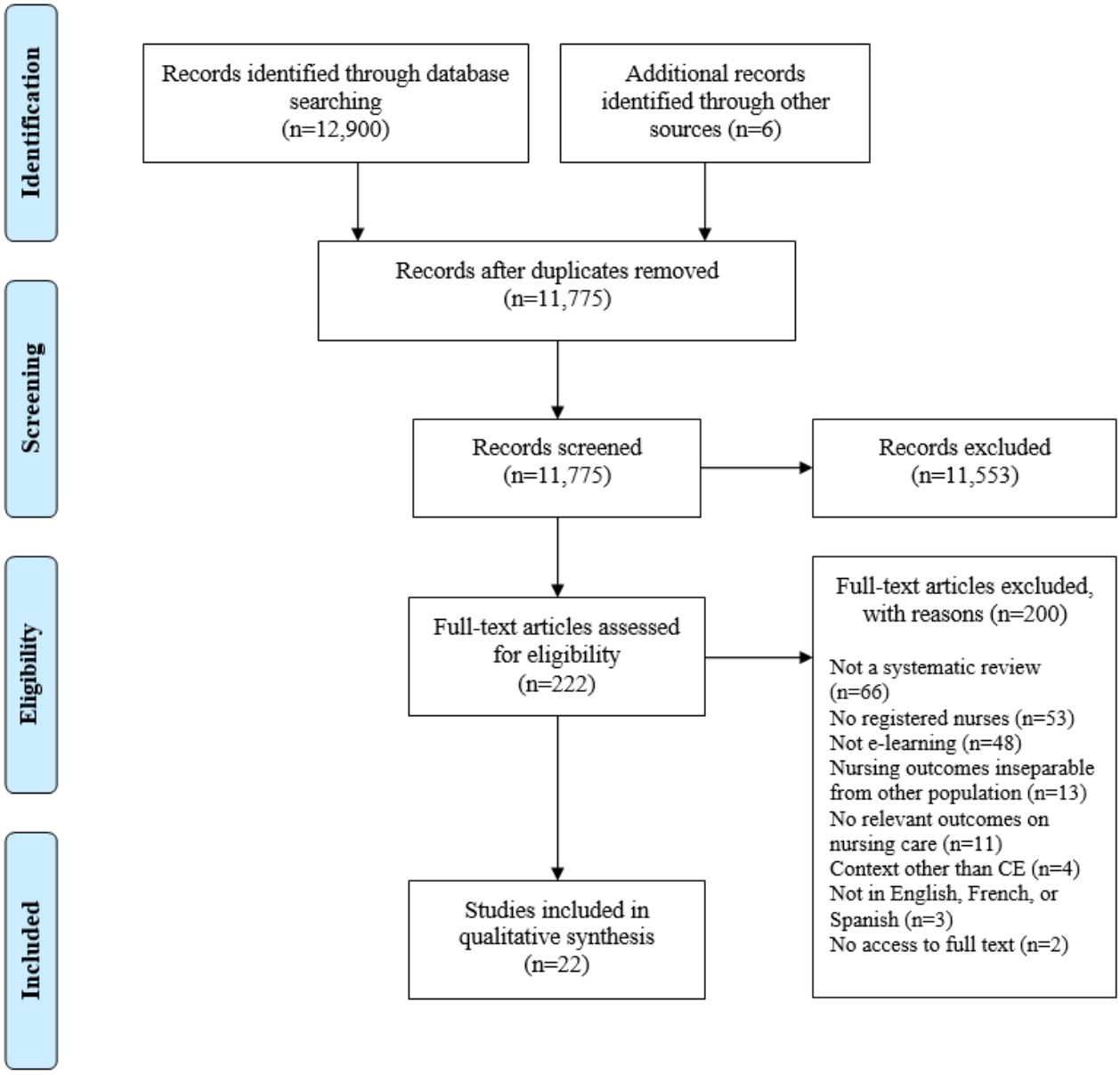
Finally, we did not calculate the corrected covered area [27] in order to measure the actual degree of overlap in the SRSRs. We simply illustrated the overlap in a matrix (see Multimedia Appendix 1), as explained earlier.

Results

Search Results and Eligibility of Systematic Reviews

The overall process of SR selection is illustrated with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram [35] (see Figure 2). We identified a total of 12,906 records. After removing duplicate references, we assessed 11,775 records for eligibility. We retrieved 222 full-text papers for detailed evaluation, from which 22 SRs published between 2008 and 2018 met the eligibility criteria. The list of included SRs is presented in Multimedia Appendix 2. In Multimedia Appendix 3, we provide the references of excluded papers as well as the reasons for exclusion.

Figure 2. Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flowchart. CE: continuing education.



Methodological Quality Assessment Results

We did not exclude papers based on methodological grounds, considering the scope of the SRSRs, which was not intended to inform action or decision making in terms of the most effective e-learning to impact nursing care. The assessment of methodological quality is presented individually for each SR (see Table 2) and globally (ie, all included SRs) using ROBIS (see Figure 3) and AMSTAR 2 (see Figure 4). Out of 22 SRs, 9 (41%) were at low risk of bias, 8 (36%) were at high risk of bias, and 5 (23%) had an unclear risk of bias. The assessment with AMSTAR 2 yielded the following results: out of 22 SRs, 6 (27%) had a high level of confidence, 4 (18%) had a moderate level of confidence, 10 (45%) had a low level of confidence, and 2 (9%) had a critically low level of confidence. The findings regarding the risk of bias and the level of confidence for the same SR were consistent across the two tools. For example, an SR [36] at high risk of bias according to the ROBIS tool was rated as having a low level of confidence using the AMSTAR 2.

Table 2. Methodological quality assessment for each individual SR[a] included in this study using a combination of the ROBIS[b] tool and the AMSTAR[c] 2.

Author, year (type of SR)	Risk of bias using the ROBIS tool, Phase 2[d]: Identifying concerns with the review process—the four domains of bias				Risk of bias using the ROBIS tool, Phase 3: Judging overall risk of bias in the review	Level of confidence using the AMSTAR 2: final judgment
	Study eligibility criteria	Identification and selection of studies	Data collection and study appraisal	Synthesis and findings		
Bloomfield, 2008 [36] (QT[e])	Low	High	High	High	High	Low
Brunero, 2012 [37] (MSR[f])	Low	Unclear	Unclear	High	Unclear	Low
Byrne, 2008 [38] (QT)	High	Unclear	High	High	Unclear	Low
Carroll, 2009 [39] (MSR)	High	Unclear	High	High	High	Critically low
Chipps, 2012 [40] (QT)	Low	Low	Low	Unclear	Low	Moderate
Coyne, 2018 [41] (MSR)	Low	Low	Low	Unclear	Unclear	Moderate
Du, 2013 [42] (QT)	Low	Low	Low	Low	Low	High
Feng, 2013 [43] (QT)	Low	Unclear	Low	Low	Low	High
Freire, 2013 [44] (MSR)	Unclear	Low	High	High	High	Low
Harkanen, 2016 [45] (QT)	Low	Low	Low	Low	Low	High
Hegland, 2017 [46] (QT)	Low	Low	Low	Low	Low	High
Hines, 2015 [47] (QT)	Low	Low	Unclear	Low	Low	Moderate
Kakushi, 2016 [48] (MSR)	Unclear	Unclear	High	High	High	Low
Kang, 2017 [49] (QT)	Low	Unclear	Low	Low	Low	High
Knapp, 2008 [10] (MSR)	High	High	High	High	High	Critically low
Lahti, 2014 [1] (QT)	Low	Unclear	Low	Low	Low	Moderate
Lam-Antoniades, 2009 [50] (QT)	Low	Unclear	Unclear	High	High	Low
Lawn, 2017 [51] (MSR)	Unclear	High	Low	High	Unclear	Low
Nicoll, 2018 [3] (MSR)	Low	Unclear	Unclear	High	High	Low
Philips, 2012 [52] (MSR)	Unclear	Unclear	Unclear	Unclear	Unclear	Low
Sinclair, 2016 [4] (QT)	Low	Low	Low	Low	Low	High
Tomlinson, 2013 [53] (QT)	Low	Unclear	High	High	High	Low

[a]SR: systematic review.

[b]ROBIS: Risk Of Bias In Systematic Reviews.

[c]AMSTAR: Assessment of Multiple Systematic Reviews.

[d]Phase 1 is optional and consists of assessing the relevance of SRs. It was not performed nor described.

[e]QT: quantitative review.

[f]MSR: mixed-studies review.

Figure 3. Methodological quality using the Risk Of Bias In Systematic Reviews (ROBIS) tool. The total risk of bias and the four domains of bias are shown. The numbers within the bars represent the number of systematic reviews.

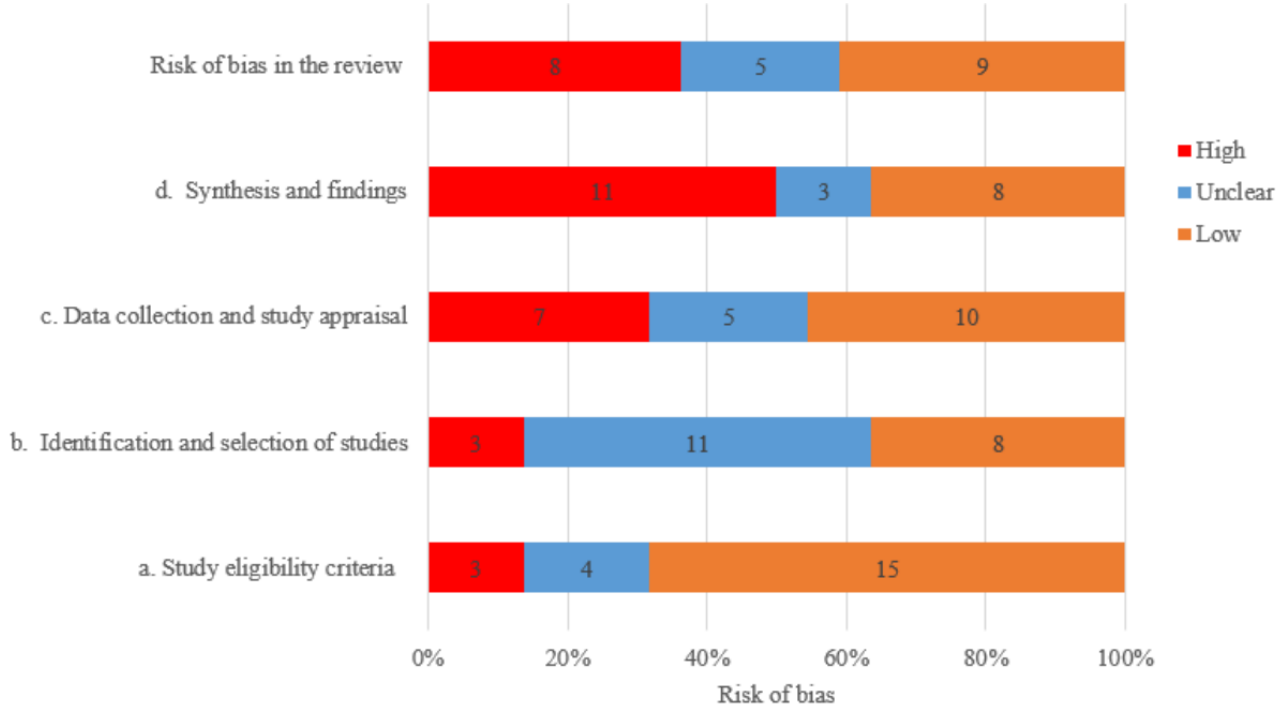
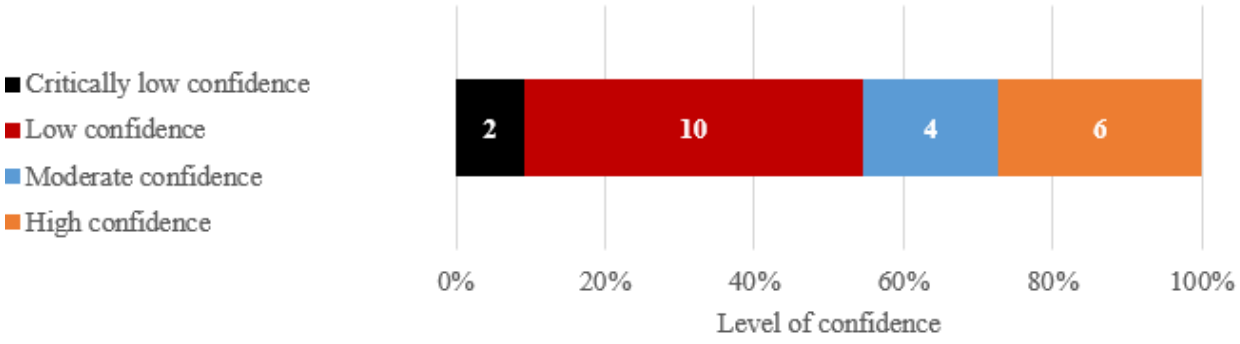


Figure 4. Methodological quality using the Assessment of Multiple Systematic Reviews (AMSTAR) 2. The numbers within the bars represent the number of systematic reviews.



General Characteristics of Systematic Reviews and Participants

General characteristics of included SRs are shown in Multimedia Appendix 4. The first authors of the included SRs were from various countries: Australia (n=7), United Kingdom (n=4), Brazil (n=2), South Africa (n=1), China (n=1), Taiwan (n=1), Korea (n=1), Finland (n=2), Norway (n=1), United States (n=1), and Canada (n=1). In 8 SRs out of 22 (36%), there was an overlap regarding primary studies (see Multimedia Appendix 1).

We included any SRs that contained one or many primary studies focusing on RNs using e-learning interventions in a CE context, which means that SRs with populations other than RNs (eg, nursing students and other health care providers) were included as long as information about RNs was clearly retrievable. The ratio of primary studies targeting nurses in a CE context to the total number of primary studies pertaining to an SR was very low. For example, from Brunero et al [37], we extracted only 2 out of the 25 (8%) primary studies that met all of the eligibility criteria. Only 1 SR [10] included all primary studies (n=5) that concerned the population of interest (ie, RNs), the e-learning intervention, the outcomes of interest, and the CE context. When reported, the number of RNs across the SRs varied from 15 [53] to 658 [3]. RNs had different job titles (eg, nurse specialists, practice nurses, community nurses, and school nurses) and worked in different settings (eg, intensive care units, emergency departments, coronary critical care, medical-surgical, pediatrics, mental health, palliative care, geriatric hospitals, and primary care).

E-Learning Interventions and Comparison Groups

There were a variety of e-learning interventions targeting nurses in a CE context (see Multimedia Appendix 5). Some examples are an online learning module regarding the use of brief motivational interviewing as a communication style to influence health behavior change [51] and online and interactive CD-ROM programs on medication administration skills and safety [45]. Other e-learning interventions were presented in terms of configuration, such as computer-assisted instructions [36], computer-based simulation [38], videoconferencing [53,52], and situated e-learning [43], while few had details on the instructional method, such as case-based learning [37].

Examples of comparison interventions included the following: electronic intervention, face-to-face intervention, no intervention, and blended learning. In four SRs [3,40,47,50], information about the theories or models used was reported regarding the engagement of stakeholders and the development or evaluation of e-learning interventions. These theories and models included engagement models [54,55], adult learning theory [56], the Kirkpatrick model [30], and the effects of information systems quality on nurses' acceptance of the e-learning system [57].

Effects of E-Learning

Overview

The outcomes are presented under different formats. First, the findings are grouped per systematic review along with a description of the interventions and the comparisons (see Multimedia Appendix 5). Second, they are described under a frequency table (see Table 3). Overall, positive outcomes (ie, effects reported in favor of e-learning interventions) are overrepresented compared to negative outcomes and those with no effect. Finally, the outcomes are synthesized and presented narratively under four main themes, informed by the Kirkpatrick model [30].

Table 3. Frequency and direction of outcomes.

Levels of evaluation from the Kirkpatrick model and subthemes	Number of documented outcomes from primary studies and direction of the effect			
	Negative	No effect	Positive	Total
1. Nurse reactions with e-learning (n=11 SRs[a])				
Total	7	0	27	34
General	0	0	9	9
Anonymity	0	0	1	1
Authentic scenario	0	0	1	1
Computer and internet experience	1	0	0	1
Confidence in e-learning	0	0	1	1
Content	0	0	1	1
Discussion	0	0	1	1
Information sharing	1	0	0	1
Interactions	2	0	2	4
Learners' experience	0	0	1	1
Overall satisfaction	0	0	1	1
Person-centered approach	0	0	1	1
Satisfaction with interactive case studies	0	0	2	2
Scope of reflection	0	0	1	1
Sense of belonging	0	0	1	1
Technical support	0	0	1	1
Technology characteristics	3	0	3	6
2. Nurse learning (n=18 SRs)				
Total	3	10	40	53
Knowledge (n=13 SRs)				
Total	1	5	18	24
General	0	3	6	9
Acute Physiology and Chronic Health Evaluation III scoring system	0	0	1	1
Arterial blood gas interpretation	0	0	1	1
Assessment (ability of neurological function)	0	0	1	1
Assessment (general)	0	0	1	1
Emergency preparedness	0	0	1	1
Hospital quality	0	0	1	1
Intravenous injections	0	0	2	2
Medication administration	0	0	1	1
Medication calculation	1	0	1	2
Neonatal care	0	0	1	1
Pain, physical and psychological symptoms, and loss	0	1	0	1
Palliative care	0	1	1	2
Attitude and self-efficacy (n=3 SRs)				
Total	0	0	4	4
Confidence postintervention	0	0	1	1
Perceived effectiveness of e-learning	0	0	1	1

Levels of evaluation from the Kirkpatrick model and subthemes	Number of documented outcomes from primary studies and direction of the effect			
	Negative	No effect	Positive	Total
Personal and professional development	0	0	1	1
Stress in nurse-patient relationship	0	0	1	1
Self-efficacy (general)	0	0	1	1
Skills (n=10 SRs)				
Total	2	5	18	25
General	0	0	4	4
Assessment (depression)	0	0	1	1
Cannulation	1	0	0	1
Cardiopulmonary resuscitation-defibrillation	1	0	0	1
Care practice changes	0	0	1	1
Child abuse detection	0	0	1	1
Communication	0	0	1	1
Critical appraisal of research literature	0	1	1	2
Emergency preparedness skills performance	0	0	1	1
Intravenous injections	0	1	0	1
Medication preparation and administration	0	3	3	6
Motivational interviewing	0	0	1	1
Monitoring	0	0	1	1
Neonatal care	0	0	1	1
Universal precautions-related behaviors	0	0	1	1
Scheduling activities	0	0	1	1
3. Behavior (change in practice) (n=0 SRs)	0	0	0	0
4. Results (n=2 SRs)				
Total	0	0	1	1
Patient outcomes (n=1 SR)				
Nurses' perceptions of care for older adults	0	0	1	1
Cost (n=2 SRs)	0	0	2	2
Total	10	10	70	90

[a]SR: systematic review.

Level 1: Nurse Reactions With E-Learning Interventions

Nurse reactions with e-learning interventions have been described in 11 of the 22 SRs (50%) [3,10,36,39-42,44,50,51,52].

Positive outcomes were described in 8 out of 22 SRs (36%) [10,36,39,40,42,44,48,51], mostly in terms of nurse satisfaction with using e-learning for the following reasons: quality of content [44], importance of social interactions [40,48], active learning [48], flexibility [10,51], effectiveness and convenience of the technology, as well as quality of support received [10]. Other sources of nurse satisfaction were reported as follows: patient-centered approach, time-saving, and self-directed learning. Nurses stressed the importance of authentic scenarios and of practicing skills in the work context [51]. Nurses found higher satisfaction with e-learning than from videotaped courses [42], while in the SR by Lam-Antoniades et al [50], nurses found that there were advantages of e-continuing education over lecture courses. Otherwise, nurses felt satisfied with both e-learning programs and traditional in-classroom programs [3].

In 3 out of 22 SRs (14%), nurse dissatisfaction with e-learning interventions was explained by the following reasons: technical difficulties [10,40], a lack of computer experience and internet literacy, slower information exchange [10], and a preference for face-to-face format [52]. In one SR [51], nurses identified access, navigation, and time as challenges.

Level 2: Nurse Learning

Overview

Nurse learning outcomes were reported in 18 of 22 SRs (82%) [1,3,4,10,36-38,40-47,49,51]. We divided learning into three subthemes: knowledge, attitude and self-efficacy, and skills.

Knowledge

In 13 SRs out of 22 (59%) [1,3,10,37,53,40-45,49,52], nurses improved their knowledge with the help of e-learning interventions on many topics, including assessment of ability of neurological function [42], medication administration and calculation [45], physiology and chronic health evaluation [10], arterial blood gas interpretation, intervention focusing on a rare disease [3], and palliative care [52]. With the help of e-learning, nurses improved their knowledge compared to no intervention [43]. However, nurse acquisition of knowledge in a classroom was superior to e-learning for drug dose calculations [45].

In 7 SRs out of 22 (32%) [10,36,53,40,46,49,52], no effects were reported on nurse knowledge. There were nonsignificant differences in knowledge scores on drug dose calculation between groups [46] and in learning effectiveness outcomes between the face-to-face versus videoconference formats [53,52]. The effect size difference reported was not significant in these 2 SRs [10,49]. There were no significant outcomes related to learning on the topics of intravenous (IV) injections and medication administration and preparation [36]. No significant change was found in nurse knowledge related to pain, physical and psychological symptoms, and loss [40].

Attitude and Self-Efficacy

Higher self-efficacy and performance scores were generally found among nurses using the e-learning intervention [42]. Nurses had positive attitudes toward effectiveness of online learning modules for motivational interviewing [51]. They perceived benefits of e-learning on their personal and professional development [52]. Other nurses improved their confidence in reducing stress in the nurse-patient relationship [37].

Skills

In 9 SRs out of 22 (41%) [3,4,36,37,42-44,46,47], positives outcomes were documented related to the increase of skills following nurses' participation in e-learning.

Nurses had better performance outcomes with e-learning compared to no intervention [42,43]. Nurses improved their skills after attending a 1-hour, e-learning-based, mental health education program on self-harm, demanding behavior, manipulation, and splitting and attention-seeking behavior [37]. These nurses also had positive comments regarding assessment, monitoring, communication, and interventions such as scheduling pleasant activities [37]. Furthermore, they rated items highly that were related to the extent to which training changed their care practices [37].

Nurses using e-learning interventions experienced positive outcomes related to universal precautions, IV injections, and medication administration [36,45]. The meta-analysis on computer-based simulation compared to other learning strategies showed significant effect in favor of e-learning for medication administration and preparation [46]. Nurses' perceived skills in performing, and clinical use of, brief motivational interviewing were more favorable postintervention [3]. Nurses had better emergency preparedness as a result of e-learning than with no intervention and they improved child abuse detection with e-learning compared to no intervention [4]. An increase in nurses' skills scores with e-learning related to neonatal care has been reported [44]. In terms of cognitive skills, nurses self-assessed their critical appraisal competencies positively regarding research literacy [47].

Negative outcomes were reported in 2 out of 22 SRs (9%). Cardiopulmonary resuscitation-defibrillation and defibrillation performance was worse among nurses using long-distance learning than that of the control group [42]. Nurses using the computer-based simulation to cannulate a real patient with force feedback had lower success at the first attempt.

Finally, 2 SRs out of 22 (9%) reported no effect by e-learning on skills. Nurses found no improvement in one critical appraisal competency related to research literacy: the identification of the sample [47]. *Core 2* errors related to preparation and administration of medication increased but the rate was not significant, as underlined in Bloomfield et al's SR [36].

Level 3: Behavior

No outcomes related to nurses' changes in practice were reported.

Level 4: Results

Patient Outcomes

In 1 of 22 SRs (5%), a positive outcome related to nurses' perceptions of care outcomes for older adults was reported [37].

Cost

In 2 SRs out of 22 (9%) [10,37], positive outcomes were reported in terms of using intranet- and CD-ROM-based education as a low-cost method of providing education for nursing staff.

Discussion

Principal Findings

Our SRSRs aimed at synthesizing qualitative and quantitative evidence regarding the effects of e-learning interventions on nursing care in a CE context. To the best of our knowledge, this is the first broad synthesis on the impact of e-learning on nurses in a CE context. As we expected, heterogeneity was found between populations (ie, RNs and workplace settings), interventions, comparisons, outcomes, types of SRs, and corresponding evidence. Conducting a meta-analysis was not the purpose of this SRSRs.

Main Outcomes: Four Levels of Evaluation

The most reported outcomes were learning (18/22, 82%), corresponding to Kirkpatrick's evaluation level 2. Nurse skills were the most frequently reported, followed by knowledge. Outcomes related to evaluation level 1 (ie, nurses' reactions with e-learning) were found in 11 out of the 22 SRs (50%). Authors of SRs described these reactions mainly with respect to technology characteristics, including perceived advantages and disadvantages (eg, navigability, technical difficulties, access, and flexibility). We found no SRs that reported outcomes regarding the translation of the content of e-learning interventions into nurses' practice and behavior (ie, evaluation level 3). This finding does not mean that e-learning had no

outcomes on practice. During the data analysis and interpretation, we used a conservative approach to classify the outcomes. Limited granularity of reported details is a well-known issue for authors of SRSRs and was observed in the included SRs. Therefore, it was difficult to know if skills, for example, improved nurses' knowledge of medication administration and preparation or if it changed nurses' practice. Only 1 SR included nurses' perceptions of patients' outcomes (ie, evaluation level 4) regarding care of elders; it also included 2 outcomes about costs. Overall, most reported outcomes were positive (n=70) as compared to negative (n=10) and neutral ones (n=10). This could indicate the presence of a reporting bias at the level of primary studies and SRs because of the disproportionate number of positive results [58].

Our findings related to the overrepresentation of the effects of e-learning interventions on reactions and learning, as well as the underrepresentation on practice and patient outcomes, are similar to those found in the literature among health care students, including nursing students [7,59]; physicians [9,60,61]; allied health practitioners [62]; and various health care providers [2]. However, Militello et al [34] conducted an SRSRs on the efficacy of computer-mediated continuing education for health care providers, including nurses, and they performed a meta-analysis. They classified their outcomes according to the Kirkpatrick model [30]. They found that 8 of the 11 SRs included measures of learner satisfaction (Level 1), 10 SRs included learning outcomes (Level 2), 9 included outcomes on provider behavior or performance (Level 3), and 5 included health and patient outcomes (Level 4). We can suppose that Militello et al [34] were more inclusive in their way of classifying outcomes related to practice change than we were in our data analysis and synthesis. Furthermore, many authors (eg, Légaré et al [9], Légaré et al [63], and Kitto et al [64]) are interested by this transition from Level 2 to Level 3 that can occur as a result of changes promoted by the content and format of continuing professional development activities, as well as how competencies are acquired and assessed. This transition not only depends on the acquisition of knowledge and skills, but also on a myriad of other elements related, for instance, to the intervention (eg, relative advantage), the outer context (eg, resources), the inner context (eg, organizational culture), individual characteristics (eg, learning style), and process (eg, planning) [65].

Methodological Quality

The methodological quality of SRs varied greatly: 59% of SRs (13/22) had an overall high or unclear risk of bias, while 55% (12/22) had a low or critically low level of confidence. Only 41% (9/22) of SRs were assessed with low risk of bias while 45% (12/22) had a moderate or high level of confidence. Our results are different from those of Militello et al [34], who synthesized the methodological quality of SRs (n=11) on computer-mediated CE for health care providers. They used 11 items from the AMSTAR [66]. Out of 11 SRs, 5 were of moderate quality and 6 were of high quality. The authors only included quantitative SRs and meta-analyses.

These findings might be explained by several reasons. We used two tools that have been designed to assess systematic quantitative reviews. When we started this SRSRs in 2017 [11], no tool was available to appraise the quality of qualitative and mixed-studies reviews. Some criteria from the ROBIS tool and the AMSTAR 2, as well as their corresponding vocabulary (eg, meta-analysis, heterogeneity, and risk of bias), were not adapted to fit with the specificities of qualitative and mixed-studies reviews. Furthermore, the systematic methodology of some included SRs was not obvious. Some authors (eg, Knapp and Byers [10] and Carroll et al [39]) mentioned the word *systematic* in their paper but they did not provide all the details to fully explain the systematic nature of their work. It is important to highlight that methodological quality is one of the three dimensions of *quality* [67]. However, methodological quality is only one dimension of critical appraisal that could be performed and it is centered on how the SR is conducted [67]. It does not capture other concepts, such as the social relevance of findings and the applicability and transferability of findings to other contexts. The dimension of conceptual clarity can also be appraised and it is related to insightfulness, including the clarity, richness, and depth of description of a phenomenon [68]. Campbell et al [69] observed that methodological and conceptual quality can be inversely correlated. It means that papers that are appraised with a low methodological quality score are usually those providing good conceptual insight. This can be partially explained by the inadequacy regarding the reporting of qualitative research methods. We recommend appreciating the richness of our findings as a means to get a broad picture of the effects of e-learning interventions on nursing care. However, our results must be interpreted with caution and are not meant to guide or inform practice, nor are they meant to determine which e-learning interventions are better in supporting CE for nurses.

Strengths and Potential Biases in the Systematic Review of Systematic Reviews Process

We used a comprehensive and systematic process throughout all stages of this SRSRs. In the search strategy, we used general keywords to explore the e-learning concept as an umbrella term, such as *virtual learning environment*, *distance learning*, *Web-based learning*, *e-learning*, and *m-learning*, among others. However, we did not use all specific key terms representing all forms of digital education, such as *serious games and gamification interventions*, *massive open online course*, *virtual reality*, and *virtual patient* [70]. Recent publications focused, for example, on serious games [59,71] and virtual reality [72], either in a context of preregistration training in health students or postregistration training among health care professionals such as nurses. Nonetheless, the use of general key terms allowed for the coverage of a wide range of potentially relevant references, considering the initial 12,906 records screened.

During the screening of titles, abstracts, and full texts, we observed that information regarding the population was sometimes misleading or incomplete, such as a population of "health students" (eg, Coyne et al [41]). In that case, instead of presuming that this abstract was not eligible based on the population, we decided to retrieve the full text. We discovered that nurses were targeted in some of these papers. Even if we were inclusive during the screening process, we may have excluded some references based on limited information provided

in titles and/or abstracts. In order to limit the risk of excluding potentially relevant papers, we conducted the screening process as a team of three reviewers.

Future Research

Our SRSRs targeted specific questions about the effects of e-learning interventions on nursing care in a CE context. Few details were provided regarding RN characteristics (eg, age and educational background) and interventions, including the SRs' instructional designs. The lack of information granularity provided by the authors of SRs [69,70] is a limitation of conducting SRSRs. Cook [73] argued that these instructional designs can have an impact on the outcomes. Furthermore, few theoretical cues were given about active ingredients pertaining to the interventions that predict or explain professional or behavior change.

We would recommend using other types of knowledge synthesis to explore complementary and broader research questions. The following are some examples:

1. What are the contexts and mechanisms through which nurses and nursing students translate knowledge and skills from e-learning interventions to their practice and, consequently, how could they lead to specific outcomes among patients? How does it work? In that case, a realist review could be performed in a digital-based nursing education and CE context, with a lens similar to the work conducted by Wong et al [74].
2. How do nurses experience e-learning interventions in their work setting? How do they describe their impact on their practice or in their environment? A meta-synthesis of qualitative studies could be done to answer these questions.

It would be useful if authors of primary studies provided enough information regarding the intervention, the context, and mechanisms, including theoretical underpinnings, which could allow researchers to understand the components that can affect outcomes.

We would also suggest exploring other types of outcomes that can be related to having e-learning interventions in workplace settings. We are in agreement with Bernt et al [62], in that the relationship between access to continuing professional development and workforce retention is unknown. Other works could be done to investigate the influence of e-learning on nursing resources or structures [19], for instance, on nurse retention and working conditions.

Furthermore, most outcomes found in the literature focus on reactions and nurses' satisfaction, learning, and change in practice. Change in knowledge and learning can be seen under a cognitivist learning approach. This approach targets the work of single individuals versus, for example, the social interactions that contribute to the learning experience of learners, seen under a social constructivist lens [3]. We would benefit from using a diversity of theoretical underpinnings, educational learning theories [75], and critical [76,77] and complexity theories [78] that have the potential to shed light on many perspectives (eg, individual, interpersonal, organizational, and sociopolitical) of envisioning education, professional development, and learners' experience.

Conclusions

The findings of this SRSRs show that the effects of e-learning are mainly reported in terms of reactions, knowledge, attitude, self-efficacy, and skills (ie, the first two evaluation levels from the Kirkpatrick model). The effectiveness of e-learning interventions used by nurses in a CE context remain unknown regarding how the learning can be transferred to change practice and affect patient outcomes. Further scientific, methodological, theoretical, and practice-based breakthroughs must feed the fast-growing field of e-learning in nursing education, especially in a life-learning perspective.

Acknowledgments

This work was supported by the Quebec Network on Nursing Intervention Research with funds from the Fonds de recherche du Québec Santé (FRQS)—Ministère de la Santé et des Services sociaux (MSSS) (grant number: 26674). GR received funding for her doctoral studies from the Canadian Institutes of Health Research (CIHR), the FRQS, the Quebec Strategy for Patient-Oriented Research—Support for People and Patient-Oriented Research and Trials (SPOR-SUPPORT) Unit, and the Quebec Network on Nursing Intervention Research.

Special thanks to Tripti Pande, who performed the search strategy, and Frédéric Bergeron, health information specialist at Université Laval, who gave valuable advice in developing the search strategy. Thank you to Gisèle Irène Claudine Mmemba who updated the search strategy.

Authors' Contributions

GR conceived and designed the SRSRs with input from MPG and JC. GR informed the search strategy. GR, JPG, and EH were responsible for data extraction. GR and JBP assessed the methodological quality of the SRs. JPG and EH were involved in the interpretation of results. GR, MPG, JC, JPG, EH, CAD, and JBP were engaged in the drafting of this manuscript and they all read and approved the final version.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Overlap between primary studies within included systematic reviews.
[XLSX File (Microsoft Excel File)19 KB - jmir_v21i10e15118_app1.xlsx]

Multimedia Appendix 2
List of included systematic reviews.
[PDF File (Adobe PDF File)77 KB - jmir_v21i10e15118_app2.pdf]

Multimedia Appendix 3
List of excluded papers.
[PDF File (Adobe PDF File)329 KB - jmir_v21i10e15118_app3.pdf]

Multimedia Appendix 4
General characteristics of included systematic reviews.
[PDF File (Adobe PDF File)325 KB - jmir_v21i10e15118_app4.pdf]

Multimedia Appendix 5
Interventions, comparisons, and outcomes from the studies.
[PDF File (Adobe PDF File)327 KB - jmir_v21i10e15118_app5.pdf]

References

1. Lahti M, Hätönen H, Välimäki M. Impact of e-learning on nurses' and student nurses knowledge, skills, and satisfaction: A systematic review and meta-analysis. Int J Nurs Stud 2014 Jan;51(1):136-149. [doi: 10.1016/j.ijnurstu.2012.12.017] [Medline: 23384695]
2. Vaona A, Banzi R, Kwag KH, Rigon G, Cereda D, Pecoraro V, et al. E-learning for health professionals. Cochrane Database Syst Rev 2018 Jan 21;1:CD011736 [FREE Full text] [doi: 10.1002/14651858.CD011736.pub2] [Medline: 29355907]
3. Nicoll P, MacRury S, van Woerden HC, Smyth K. Evaluation of technology-enhanced learning programs for health care professionals: Systematic review. J Med Internet Res 2018 Apr 11;20(4):e131 [FREE Full text] [doi: 10.2196/jmir.9085] [Medline: 29643049]
4. Sinclair PM, Kable A, Levett-Jones T, Booth D. The effectiveness of Internet-based e-learning on clinician behaviour and patient outcomes: A systematic review. Int J Nurs Stud 2016 May;57:70-81. [doi: 10.1016/j.ijnurstu.2016.01.011] [Medline: 27045566]
5. De Caro W, Marucci AR, Lancia L, Sansoni J. Case study in nursing. In: Biondi-Zoccai G, editor. Umbrella Reviews: Evidence Synthesis with Overviews of Reviews and Meta-Epidemiologic Studies. Cham, Switzerland: Springer International Publishing; 2016:273-303.
6. Clark RC, Mayer RE. E-Learning and the Science of Instruction: Proven Guidelines for Consumers and Designers of Multimedia Learning. 4th edition. Hoboken, NJ: John Wiley & Sons; 2016.
7. Voutilainen A, Saaranen T, Sormunen M. Conventional vs e-learning in nursing education: A systematic review and meta-analysis. Nurse Educ Today 2017 Mar;50:97-103. [doi: 10.1016/j.nedt.2016.12.020] [Medline: 28038371]
8. Canadian Nurses Association. Framework for the Practice of Registered Nurses in Canada. Ottawa, ON: Canadian Nurses Association; 2015. URL: https://www.cna-aiic.ca/~/media/cna/page-content/pdf-en/framework-for-the-pracice-of-registered-nurses-in-canada.pdf?la=en [accessed 2019-06-06]
9. Légaré F, Freitas A, Thompson-Leduc P, Borduas F, Luconi F, Boucher A, et al. The majority of accredited continuing professional development activities do not target clinical behavior change. Acad Med 2015 Feb;90(2):197-202. [doi: 10.1097/ACM.0000000000000543] [Medline: 25354076]
10. Knapp SJ, Byers JF. Use of the Internet in staff development and its application in helping critical care nurses to lower family stress. J Nurses Staff Dev 2008;24(1):E1-E8. [doi: 10.1097/01.NND.0000300860.37126.37] [Medline: 18349757]
11. Rouleau G, Gagnon MP, Côté J, Payne-Gagnon J, Hudson E, Bouix-Picasso J, et al. Effects of e-learning in a continuing education context on nursing care: A review of systematic qualitative, quantitative and mixed studies reviews (protocol). BMJ Open 2017 Oct 16;7(10):e018441 [FREE Full text] [doi: 10.1136/bmjopen-2017-018441] [Medline: 29042394]
12. De Caro W, Marucci AR, Giordani M, Sansoni J. E-learning and university nursing education: An overview of reviews [Article in Italian]. Prof Inferm 2014;67(2):107-116. [doi: 10.7429/pi.2013.672107] [Medline: 25134508]
13. Cook DA, Levinson AJ, Garside S, Dupras DM, Erwin PJ, Montori VM. Internet-based learning in the health professions: A meta-analysis. JAMA 2008 Sep 10;300(10):1181-1196. [doi: 10.1001/jama.300.10.1181] [Medline: 18780847]
14. Hunt H, Pollock A, Campbell P, Estcourt L, Brunton G. An introduction to overviews of reviews: Planning a relevant research question and objective for an overview. Syst Rev 2018 Mar 01;7(1):39 [FREE Full text] [doi: 10.1186/s13643-018-0695-8] [Medline: 29490699]

15. O'Connor D, Green S, Higgins JPT. Cochrane Handbook for Systematic Reviews of Interventions. Version 5.1.0 (updated March 2011). London, UK: The Cochrane Collaboration; 2011. Defining the review question and developing criteria for including studies URL: https://handbook-5-1.cochrane.org/chapter_5/5_defining_the_review_question_and_developing_criteria_for.htm [accessed 2019-05-07]
16. Pollock M, Fernandes RM, Becker LA, Featherstone R, Hartling L. What guidance is available for researchers conducting overviews of reviews of healthcare interventions? A scoping review and qualitative metasummary. Syst Rev 2016 Nov 14;5(1):190 [FREE Full text] [doi: 10.1186/s13643-016-0367-5] [Medline: 27842604]
17. Pollock M, Fernandes R, Becker L, Pieper D, Hartling L. Chapter V: Overviews of reviews. Draft version (October 8, 2018). In: Cochrane Handbook for Systematic Reviews of Interventions. London, UK: Cochrane Collaboration; 2018.
18. Henriksen K, Dayton E, Keyes M, Carayon P, Hughes R. Patient Safety and Quality: An Evidence-Based Handbook for Nurses. Rockville, MD: AHRQ Publication; 2008. Understanding adverse events: A human factors framework URL: https://www.umbrellacaremanagement.com/wp-content/uploads/2018/02/nurseshdbk.pdf [accessed 2019-07-30]
19. Dubois CA, D'Amour D, Pomey M, Girard F, Brault I. Conceptualizing performance of nursing care as a prerequisite for better measurement: A systematic and interpretive review. BMC Nurs 2013 Mar 07;12:7 [FREE Full text] [doi: 10.1186/1472-6955-12-7] [Medline: 23496961]
20. Rouleau G, Gagnon MP, Côté J, Payne-Gagnon J, Hudson E, Dubois CA. Impact of information and communication technologies on nursing care: Results of an overview of systematic reviews. J Med Internet Res 2017 Apr 25;19(4):e122 [FREE Full text] [doi: 10.2196/jmir.6686] [Medline: 28442454]
21. Systematic Reviews: CRD's Guidance for Undertaking Reviews in Health Care. York, UK: Centre for Reviews and Dissemination, University of York; 2009 Jan. URL: https://www.york.ac.uk/media/crd/Systematic_Reviews.pdf [accessed 2017-05-07]
22. Shea BJ, Reeves BC, Wells G, Thuku M, Hamel C, Moran J, et al. AMSTAR 2: A critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. BMJ 2017 Sep 21;358:j4008 [FREE Full text] [doi: 10.1136/bmj.j4008] [Medline: 28935701]
23. Whiting P, Savović J, Higgins JPT, Caldwell DM, Reeves BC, Shea B, ROBIS group. ROBIS: A new tool to assess risk of bias in systematic reviews was developed. J Clin Epidemiol 2016 Jan;69:225-234 [FREE Full text] [doi: 10.1016/j.jclinepi.2015.06.005] [Medline: 26092286]
24. Higgins JPT, Lasserson T, Chandler J, Tovey D, Churchill R. Methodological Expectations of Cochrane Intervention Reviews. London, UK: Cochrane; 2019 Jul. Standards for the conduct of new Cochrane Intervention Reviews URL: https://community.cochrane.org/book_pdf/550 [accessed 2019-07-24]
25. Pluye P, Hong QN. Combining the power of stories and the power of numbers: Mixed methods research and mixed studies reviews. Annu Rev Public Health 2014;35:29-45. [doi: 10.1146/annurev-publhealth-032013-182440] [Medline: 24188053]
26. Creswell JW. Research Design: Qualitative, Quantitative, and Mixed-Methods Approaches. 4th edition. Thousand Oaks, CA: SAGE Publications; 2014.
27. Pieper D, Antoine S, Mathes T, Neugebauer EA, Eikermann M. Systematic review finds overlapping reviews were not mentioned in every other overview. J Clin Epidemiol 2014 Apr;67(4):368-375. [doi: 10.1016/j.jclinepi.2013.11.007] [Medline: 24581293]
28. Pollock M, Fernandes RM, Newton AS, Scott SD, Hartling L. A decision tool to help researchers make decisions about including systematic reviews in overviews of reviews of healthcare interventions. Syst Rev 2019 Jan 22;8(1):29 [FREE Full text] [doi: 10.1186/s13643-018-0768-8] [Medline: 30670086]
29. Pollock M, Fernandes RM, Newton AS, Scott SD, Hartling L. The impact of different inclusion decisions on the comprehensiveness and complexity of overviews of reviews of healthcare interventions. Syst Rev 2019 Jan 11;8(1):18 [FREE Full text] [doi: 10.1186/s13643-018-0914-3] [Medline: 30635048]
30. Kirkpatrick DL, Kirkpatrick JD. Evaluating Training Programs: The Four Levels. 3rd edition. San Francisco, CA: Berrett-Koehler Publishers; 2006.
31. Shen N, Yufe S, Saadatfard O, Sockalingam S, Wiljer D. Rebooting Kirkpatrick: Integrating information system theory into the evaluation of Web-based continuing professional development interventions for interprofessional education. J Contin Educ Health Prof 2017;37(2):137-146. [doi: 10.1097/CEH.0000000000000154] [Medline: 28562503]
32. Moore DE, Green JS, Gallis HA. Achieving desired results and improved outcomes: Integrating planning and assessment throughout learning activities. J Contin Educ Health Prof 2009;29(1):1-15. [doi: 10.1002/chp.20001] [Medline: 19288562]
33. Jeffries PR. A framework for designing, implementing, and evaluating simulations used as teaching strategies in nursing. Nurs Educ Perspect 2005;26(2):96-103. [Medline: 15921126]
34. Militello LK, Gance-Cleveland B, Aldrich H, Kamal R. A methodological quality synthesis of systematic reviews on computer-mediated continuing education for healthcare providers. Worldviews Evid Based Nurs 2014 Jun;11(3):177-186. [doi: 10.1111/wvn.12041] [Medline: 24865984]
35. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JP, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: Explanation and elaboration. PLoS Med 2009 Jul 21;6(7):e1000100 [FREE Full text] [doi: 10.1371/journal.pmed.1000100] [Medline: 19621070]

36. Bloomfield JG, While AE, Roberts JD. Using computer assisted learning for clinical skills education in nursing: Integrative review. J Adv Nurs 2008 Aug;63(3):222-235. [doi: 10.1111/j.1365-2648.2008.04653.x] [Medline: 18702768]
37. Brunero S, Jeon Y, Foster K. Mental health education programmes for generalist health professionals: An integrative review. Int J Ment Health Nurs 2012 Oct;21(5):428-444. [doi: 10.1111/j.1447-0349.2011.00802.x] [Medline: 22500589]
38. Byrne AJ, Pugsley L, Hashem MA. Review of comparative studies of clinical skills training. Med Teach 2008;30(8):764-767. [doi: 10.1080/01421590802279587] [Medline: 18946819]
39. Carroll C, Booth A, Papaioannou D, Sutton A, Wong R. UK health-care professionals' experience of on-line learning techniques: A systematic review of qualitative data. J Contin Educ Health Prof 2009;29(4):235-241. [doi: 10.1002/chp.20041] [Medline: 19998474]
40. Chipps J, Brysiewicz P, Mars M. A systematic review of the effectiveness of videoconference-based tele-education for medical and nursing education. Worldviews Evid Based Nurs 2012 Apr;9(2):78-87. [doi: 10.1111/j.1741-6787.2012.00241.x] [Medline: 22409341]
41. Coyne E, Rands H, Frommolt V, Kain V, Plugge M, Mitchell M. Investigation of blended learning video resources to teach health students clinical skills: An integrative review. Nurse Educ Today 2018 Apr;63:101-107. [doi: 10.1016/j.nedt.2018.01.021] [Medline: 29425738]
42. Du S, Liu Z, Liu S, Yin H, Xu G, Zhang H, et al. Web-based distance learning for nurse education: A systematic review. Int Nurs Rev 2013 Jun;60(2):167-177. [doi: 10.1111/inr.12015] [Medline: 23691999]
43. Feng J, Chang Y, Chang H, Erdley WS, Lin C, Chang Y. Systematic review of effectiveness of situated e-learning on medical and nursing education. Worldviews Evid Based Nurs 2013 Aug;10(3):174-183. [doi: 10.1111/wvn.12005] [Medline: 23510119]
44. Freire LM, Paula MA, Duarte ED, Bueno M. Distance education in neonatal nursing scenarios: A systematic review [Article in Portuguese]. Rev Esc Enferm USP 2015 Jun;49(3):515-521 [FREE Full text] [doi: 10.1590/S0080-623420150000300021] [Medline: 26107714]
45. Härkänen M, Voutilainen A, Turunen E, Vehviläinen-Julkunen K. Systematic review and meta-analysis of educational interventions designed to improve medication administration skills and safety of registered nurses. Nurse Educ Today 2016 Jun;41:36-43. [doi: 10.1016/j.nedt.2016.03.017] [Medline: 27138480]
46. Hegland PA, Aarlie H, Strømme H, Jamtvedt G. Simulation-based training for nurses: Systematic review and meta-analysis. Nurse Educ Today 2017 Jul;54:6-20. [doi: 10.1016/j.nedt.2017.04.004] [Medline: 28456053]
47. Hines S, Ramsbotham J, Coyer F. The effectiveness of interventions for improving the research literacy of nurses: A systematic review. Worldviews Evid Based Nurs 2015 Oct;12(5):265-272. [doi: 10.1111/wvn.12106] [Medline: 26422084]
48. Kakushi LE, Évora YD. Social networking in nursing education: Integrative literature review. Rev Lat Am Enfermagem 2016;24:e2709 [FREE Full text] [doi: 10.1590/1518-8345.1055.2709] [Medline: 27384465]
49. Kang J, Seomun G. Evaluating Web-based nursing education's effects: A systematic review and meta-analysis. West J Nurs Res 2018 Nov;40(11):1677-1697. [doi: 10.1177/0193945917729160] [Medline: 28875770]
50. Lam-Antoniades M, Ratnapalan S, Tait G. Electronic continuing education in the health professions: An update on evidence from RCTs. J Contin Educ Health Prof 2009;29(1):44-51. [doi: 10.1002/chp.20005] [Medline: 19288566]
51. Lawn S, Zhi X, Morello A. An integrative review of e-learning in the delivery of self-management support training for health professionals. BMC Med Educ 2017 Oct 10;17(1):183 [FREE Full text] [doi: 10.1186/s12909-017-1022-0] [Medline: 29017521]
52. Phillips JL, Piza M, Ingham J. Continuing professional development programmes for rural nurses involved in palliative care delivery: An integrative review. Nurse Educ Today 2012 May;32(4):385-392. [doi: 10.1016/j.nedt.2011.05.005] [Medline: 21641095]
53. Tomlinson J, Shaw T, Munro A, Johnson R, Madden DL, Phillips R, et al. How does tele-learning compare with other forms of education delivery? A systematic review of tele-learning educational outcomes for health professionals. N S W Public Health Bull 2013 Nov;24(2):70-75 [FREE Full text] [doi: 10.1071/NB12076] [Medline: 24195848]
54. Seibert DC, Guthrie JT, Adamo G. Improving learning outcomes: Integration of standardized patients & telemedicine technology. Nurs Educ Perspect 2004;25(5):232-237. [Medline: 15508562]
55. Stain SC, Mitchell M, Belue R, Mosley V, Wherry S, Adams CZ, et al. Objective assessment of videoconferenced lectures in a surgical clerkship. Am J Surg 2005 Jan;189(1):81-84. [doi: 10.1016/j.amjsurg.2004.04.012] [Medline: 15701498]
56. Conole G, Dyke M, Oliver M, Seale J. Mapping pedagogy and tools for effective learning design. Comput Educ 2004 Aug;43(1-2):17-33. [doi: 10.1016/j.compedu.2003.12.018]
57. Cheng Y. The effects of information systems quality on nurses' acceptance of the electronic learning system. J Nurs Res 2012 Mar;20(1):19-30. [doi: 10.1097/JNR.0b013e31824777aa] [Medline: 22333963]
58. Mlinarić A, Horvat M, Šupak Smolčić V. Dealing with the positive publication bias: Why you should really publish your negative results. Biochem Med (Zagreb) 2017 Oct 15;27(3):030201 [FREE Full text] [doi: 10.11613/BM.2017.030201] [Medline: 29180912]
59. Gentry SV, Gauthier A, L'Estrade Ehrstrom B, Wortley D, Lilienthal A, Tudor Car L, et al. Serious gaming and gamification education in health professions: Systematic review. J Med Internet Res 2019 Mar 28;21(3):e12994 [FREE Full text] [doi: 10.2196/12994] [Medline: 30920375]

60. Cho D, Cosimini M, Espinoza J. Podcasting in medical education: A review of the literature. Korean J Med Educ 2017 Dec;29(4):229-239 [FREE Full text] [doi: 10.3946/kjme.2017.69] [Medline: 29207454]
61. Chen F, Lui AM, Martinelli SM. A systematic review of the effectiveness of flipped classrooms in medical education. Med Educ 2017 Jun;51(6):585-597. [doi: 10.1111/medu.13272] [Medline: 28488303]
62. Berndt A, Murray CM, Kennedy K, Stanley MJ, Gilbert-Hunt S. Effectiveness of distance learning strategies for continuing professional development (CPD) for rural allied health practitioners: A systematic review. BMC Med Educ 2017 Jul 12;17(1):117 [FREE Full text] [doi: 10.1186/s12909-017-0949-5] [Medline: 28701199]
63. Légaré F, Borduas F, Freitas A, Jacques A, Godin G, Luconi F, CPD-KT team. Development of a simple 12-item theory-based instrument to assess the impact of continuing professional development on clinical behavioral intentions. PLoS One 2014;9(3):e91013 [FREE Full text] [doi: 10.1371/journal.pone.0091013] [Medline: 24643173]
64. Kitto S, Danilovich N, Delva D, Meuser J, Presseau J, Grimshaw J, et al. Uncharted territory: Knowledge translation of competency-based continuing professional development in family medicine. Can Fam Physician 2018 Apr;64(4):250-253 [FREE Full text] [Medline: 29650601]
65. Damschroder LJ, Aron DC, Keith RE, Kirsh SR, Alexander JA, Lowery JC. Fostering implementation of health services research findings into practice: A consolidated framework for advancing implementation science. Implement Sci 2009 Aug 07;4:50 [FREE Full text] [doi: 10.1186/1748-5908-4-50] [Medline: 19664226]
66. Shea BJ, Hamel C, Wells GA, Bouter LM, Kristjansson E, Grimshaw J, et al. AMSTAR is a reliable and valid measurement tool to assess the methodological quality of systematic reviews. J Clin Epidemiol 2009 Oct;62(10):1013-1020. [doi: 10.1016/j.jclinepi.2008.10.009] [Medline: 19230606]
67. Hong QN, Pluye P. A conceptual framework for critical appraisal in systematic mixed studies reviews. J Mix Methods Res 2018 Apr 21:1. [doi: 10.1177/1558689818770058]
68. Toye F, Seers K, Allcock N, Briggs M, Carr E, Andrews J, et al. 'Trying to pin down jelly': Exploring intuitive processes in quality assessment for meta-ethnography. BMC Med Res Methodol 2013 Mar 21;13:46 [FREE Full text] [doi: 10.1186/1471-2288-13-46] [Medline: 23517438]
69. Campbell R, Pound P, Morgan M, Daker-White G, Britten N, Pill R, et al. Evaluating meta-ethnography: Systematic analysis and synthesis of qualitative research. Health Technol Assess 2011 Dec;15(43):1-164 [FREE Full text] [doi: 10.3310/hta15430] [Medline: 22176717]
70. Car J, Carlstedt-Duke J, Tudor Car L, Posadzki P, Whiting P, Zary N, Digital Health Education Collaboration. Digital education in health professions: The need for overarching evidence synthesis. J Med Internet Res 2019 Feb 14;21(2):e12913 [FREE Full text] [doi: 10.2196/12913] [Medline: 30762583]
71. Ijaz A, Khan M, Ali S, Qadir J, Boulos M. Serious games for healthcare professional training: A systematic review. Eur J Biomed Inform 2019;15:12-28 [FREE Full text]
72. Kyaw BM, Saxena N, Posadzki P, Vseteckova J, Nikolaou CK, George PP, et al. Virtual reality for health professions education: Systematic review and meta-analysis by the Digital Health Education Collaboration. J Med Internet Res 2019 Jan 22;21(1):e12959 [FREE Full text] [doi: 10.2196/12959] [Medline: 30668519]
73. Cook DA, Levinson AJ, Garside S, Dupras DM, Erwin PJ, Montori VM. Instructional design variations in Internet-based learning for health professions education: A systematic review and meta-analysis. Acad Med 2010 May;85(5):909-922. [doi: 10.1097/ACM.0b013e3181d6c319] [Medline: 20520049]
74. Wong G, Greenhalgh T, Pawson R. Internet-based medical education: A realist review of what works, for whom and in what circumstances. BMC Med Educ 2010 Feb 02;10:12 [FREE Full text] [doi: 10.1186/1472-6920-10-12] [Medline: 20122253]
75. Lavoie P, Michaud C, Bélisle M, Boyer L, Gosselin E, Grondin M, et al. Learning theories and tools for the assessment of core nursing competencies in simulation: A theoretical review. J Adv Nurs 2018 Feb;74(2):239-250. [doi: 10.1111/jan.13416] [Medline: 28815750]
76. Mooney M, Nolan L. A critique of Freire's perspective on critical social theory in nursing education. Nurse Educ Today 2006 Apr;26(3):240-244. [doi: 10.1016/j.nedt.2005.10.004] [Medline: 16330133]
77. Josephsen J. Critically reflexive theory: A proposal for nursing education. Adv Nurs 2014;2014:1-7. [doi: 10.1155/2014/594360]
78. Opfer VD, Pedder D. Conceptualizing teacher professional learning. Rev Educ Res 2011 Sep;81(3):376-407. [doi: 10.3102/0034654311413609]

Abbreviations

AMSTAR: Assessment of Multiple Systematic Reviews
CE: continuing education
CIHR: Canadian Institutes of Health Research
CINAHL: Cumulative Index of Nursing and Allied Health Literature
FRQS: Fonds de recherche du Québec Santé
IV: intravenous

JBI: Joanna Briggs Institute
MSR: mixed-studies review
MSSS: Ministère de la Santé et des Services sociaux
NCPF: Nursing Care Performance Framework
PICOS: population, intervention, comparison, outcomes, and study design
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PROSPERO: International Prospective Register of Systematic Reviews
QT: quantitative review
RN: registered nurse
ROBIS: Risk Of Bias In Systematic Reviews
SPOR-SUPPORT: Strategy for Patient-Oriented Research—Support for People and Patient-Oriented Research and Trials
SR: systematic review
SRSRs: systematic review(s) of systematic reviews

Edited by G Eysenbach; submitted 24.06.19; peer-reviewed by I Paraskakis, S Sarbadhikari, E Ray Chaudhuri, B Worm, T Taveira-Gomes; comments to author 16.07.19; revised version received 30.07.19; accepted 31.07.19; published 02.10.19.

Please cite as:
Rouleau G, Gagnon MP, Côté J, Payne-Gagnon J, Hudson E, Dubois CA, Bouix-Picasso J
Effects of E-Learning in a Continuing Education Context on Nursing Care: Systematic Review of Systematic Qualitative, Quantitative, and Mixed-Studies Reviews
J Med Internet Res 2019;21(10):e15118
URL: https://www.jmir.org/2019/10/e15118
doi:10.2196/15118
PMID:31579016

Original Paper

Codifying Online Social Support for Breast Cancer Patients: Retrospective Qualitative Assessment

Jude P Mikal[1], PhD; Stuart W Grande[1], MPA, PhD; Michael J Beckstrand[2], PhD

[1]Division of Health Policy and Management, School of Public Health, University of Minnesota, Minneapolis, MN, United States

[2]College of Liberal Arts, University of Minnesota, Minneapolis, MN, United States

Corresponding Author:
Jude P Mikal, PhD
Division of Health Policy and Management
School of Public Health
University of Minnesota
Mayo Building and Additions
420 Delaware St SE
Minneapolis, MN, 55455
United States
Phone: 1 6126264182
Email: jpmikal@umn.edu

Abstract

Background: Social media has emerged as the epicenter for exchanging health-related information, resources, and emotional support. However, despite recognized benefits of social media for advancing health-promoting support exchange, researchers have struggled to differentiate between the different ways social support occurs and is expressed through social media.

Objective: The objective of this study was to develop a fuller understanding of social support exchange by examining the ways in which breast cancer patients discuss their health needs and reach out for support on Facebook and to develop a coding schema that can be useful to other social media researchers.

Methods: We conducted a retrospective qualitative assessment of text-based social support exchanges through Facebook among 30 breast cancer survivors. Facebook wall data were systematically scraped, organized, coded, and characterized by whether and which types of support were exchanged. Research questions focused on how often participants posted related to cancer, how often cancer patients reached out for support, and the relative frequency of informational, instrumental, or socioemotional support requests broadcast by patients on the site.

Results: A novel ground-up coding schema applied to unwieldy Facebook data successfully identified social support exchange in two critical transitions in cancer treatment: diagnosis and transition off cancer therapy. Explanatory coding, design, and analysis processes led to a novel coding schema informed by 100,000 lines of data, an a priori literature review, and observed online social support exchanges. A final coding schema permits a compelling analysis of support exchange as a type of peer community, where members act proactively to buffer stress effects associated with negative health experiences. The coding schema framed operational definitions of what support meant and the forms each type of support could take in social media spaces.

Conclusions: Given the importance of social media in social interaction, support exchange, and health promotion, our findings provide insight and clarity for researchers into the different forms informational, resource, and emotional support may take in Web-based social environments. Findings support broader continuity for evaluating computer-mediated support exchange.

(J Med Internet Res 2019;21(10):e12880) doi:10.2196/12880

KEYWORDS

social support; social networking; social media; health communication; breast cancer

Introduction

Background

More and more people are using social media to gather health information and access social support, and this trend has practical implications across health research and interventions. With nearly 7 in 10 adults using some type of social media, there are limitless opportunities for sharing, learning, and exchanging ideas [1,2]. Despite recognized benefits of social media for advancing lifestyle and wellness communication [3], researchers have had a much harder time differentiating between the different ways in which social support is expressed via social media. With some exceptions [4,5], research describing social support exchange via social media has struggled to provide a structure to the content and evolution of social media use following health and other transitions and trauma. Hampered by a tidal wave of not only social media data but also social media *platforms*, researchers have struggled to cast a net around a temporal swatch of data for a group of participants that is large enough to permit a systematic analysis of generalizable trends, as well as differences within and between subjects.

Researchers have used a number of research design and analysis techniques to permit a more systematic analysis of social media use corresponding to (or in response to) health transition and trauma. This includes data gathered from controlled or condition-specific sites [6-8], single snapshots of data or cross-sectional research design [9,10], or qualitative accounts of support exchange during times of health transition [11,12]. However, although each approach provides unique insights into the nature, quality, and health correlates of support exchanged via social media, the majority provide limited information on *how* social support is expressed or evolved in response to shifting support needs. Differentiating ways social support is given and received, what forms support takes on social media, and what online support means to users will advance learning beyond basic observation toward more robust analysis of this highly interactive environment.

At the outset, the goal of this study was to calculate the proportion of Facebook exchanges that included the exchange of social support. However, in codifying support exchange, a number of challenges emerged with respect to measurement and categorization of social support exchange on social media. In this paper, we present some of the unanticipated challenges associated with coding unwieldy social media data, the systematic approach our team used to manage those challenges, and the coding scheme that resulted. The goal of this paper is to lay the foundation for future studies coding, categorizing, and measuring social support exchanges via Facebook and begin to lay a foundation for analysis of how online support exchange comes to bear upon observed mental and physical health outcomes.

Theoretical Framing

Social relationships and support have consistently emerged in research as important determinants of health (for a review, see [13,14]), but as the media through which those relationships are maintained grow and develop, so must the models used to describe them, their measure, and the techniques used to evaluate their effectiveness. In this review of relevant literature, we describe the evolution of social support research from the acknowledgment of a link between social networks and health, through the advent and widespread adoption of computer-mediated communication, to the persistent lack of clear and accessible published schemas for the systematic categorization of support in Web-based communities.

Social Support and Health

As early as 1951, theorists have acknowledged the important stress buffering effects of social engagement [15]. Since then, scholars have demonstrated how participation in social organizations can create networks of support that diminish stress and ultimately bolster health [16-20]. Despite wide acknowledgment of the relationship between social relationships and health, the measurement and mechanisms through which support impacts health has been the subject of considerable discussion. For instance, whereas a study has proposed that support effectiveness is contingent upon structural aspects of the networks within which an individual is embedded [21], other study has focused on the stress buffering effects stemming from the nature and quality of the support exchanges themselves [21,22].

However, from this study, two notable continuities have emerged. The first is the distinction between the various types of support that can be exchanged. When support is exchanged, there is a general continuity in the research literature that such support will be informational, instrumental, or socioemotional. Informational support refers to information exchanged in response to an event or problem. Instrumental support refers to the exchange of resources designed to aid in coping, and socioemotional support refers to the support derived from feelings of togetherness, esteem, or belonging [23]. In addition, research has largely settled on three caveats for support effectiveness. In other words, to be effective, research has found that social support should be both empathetic [24] and responsive to the support needs expressed [22], and support should not appear burdensome to the support provider [25]. This presents important questions for support communicated via the internet that may diminish the cost of support exchange but simultaneously undermine key elements of support quality, such as fit and empathy.

Social Support and Cancer

In addition to physical challenges, cancer diagnosis brings about a host of deleterious psychological impacts that last throughout treatment and into the transition to survivorship [26-28]. However, research shows that social support can buffer against the ill-effects of stress in three ways: (1) by bolstering morale and impacting threat assessment, (2) by impacting an individual's behavior and promoting health positive behaviors, or (3) through biological processes, such as a lowered heart rate, altered hormone production, or improved immune function [13,14,29,30]. In the case of cancer patients, research has shown that strong networks of support can buffer against stress by bolstering mental and, by extension, physical health. And research shows that the implications of high or low social engagement are long-lasting. For example, research on breast cancer patients showed that lower levels of support at diagnosis

were associated with a 4-fold increase in all-cause mortality and a 2-fold increase in mortality resulting from breast cancer [31]. And on the flip side, Salonen et al [32] found that received support had a positive impact on physical health and quality of life following breast cancer surgery. Even after the transition off cancer treatment, face-to-face networks of support have been shown to play an important role in determining physical health, post traumatic growth, and psychological well-being [33,34].

Online Support-Seeking

Despite the known health benefits of positive face-to-face supportive interactions, cancer diagnosis and treatment may correspond to a time of limited support access. Patients may find that previously established networks of support are either unwilling to provide support or unable to meet the support needs that emerge throughout cancer treatment [5,35,36]. As a result, patients shift their attention to Web networks of support as possibly better suited to providing continuous access to consistent and fitting support [5]. With more than three-quarters of Americans accessing the internet between *daily* and *near constantly* [37], Web-based social relationships have become ubiquitous, and the internet has become an important tool in accessing not only health information but also dynamic networks of individuals coping with similar conditions. In addition, research shows that the internet is an effective medium for support exchange and offers a host of functional advantages over face-to-face support, including better fitting support [38], lower barriers to accessing support [39], increased control and privacy [40], and reduced reciprocal obligation [41].

Internet-Mediated Social Support and Cancer

Research on support for breast cancer patients specifically has shown that computer-mediated communication is an effective medium for support transmission spanning from early diagnosis to treatment and into cancer survivorship. Early research on internet support groups for patients with breast cancer found that Web-based groups confer a host of mental and physical health benefits, including the following: improvement in quality of life, psychological symptoms, and coping response as well as reduction in pain [42-45]. Although trial evidence is minimal and limited, observational data strongly support the value of breast cancer support groups for quality of life and reduced feelings of depression and anxiety [7,46]. In a qualitative research study of 15 breast cancer patients, Hoybye et al [12] showed that social support communicated on the Web promoted personal empowerment and the exchange of knowledge between breast cancer patients and survivors. These results echo findings from a quantitative study of 206 breast cancer patients that showed that internet-mediated social support was associated with an increase in knowledge about cancer and its treatment options, along with a decrease in overall anxiety in patients [47].

However, with respect to Facebook—a multifaceted platform that provides access to varied networks of support—research has typically focused on support exchanged in contexts characterized by the same general structure: forum-style Facebook groups. Research results from this area of study have been mixed. Whereas a study shows that engagement with breast cancer support groups has the potential to promote empathy and the formation of personal relationships based on shared experiences [48], other studies suggest that the benefits of such groups can be exclusive, and the benefits of engagement are not distributed evenly among participants [49,50]. Nevertheless, the considerable variation within Facebook with respect to the media, and varied mechanisms of support exchange can make it difficult to differentiate types of behaviors observed on the Web and link those patterns with health benefits [1]. So, although Facebook groups are an important conduit of informational, esteem, companionship, or resource support during times of stress, they represent only one aspect of support exchange via Facebook. In this study, we aimed to create the building blocks to permit quantification, evaluation, and analysis of support exchange between individuals that could be used outside the context of Facebook groups.

Methods

Overview

The coding scheme developed by our team was designed to help organize unwieldy social media data, allow us to evaluate patterns in social media use, and measure changes in those patterns following transitions in cancer care. Methodical coding required a series of definitions and decisions based on previous literature, observational data, and bottom-up contextual analysis and those definitions are often edited out of papers and manuscripts because of space constraints. The purpose of this paper is to describe how the data created ambiguity in defining support and our rationale for decision making and operationalize and standardize the measurement of support communicated via social media, so that other researchers can adopt and adapt as needed.

To qualitatively assess the types of support exchanged through social media along with the challenges associated with categorization of these text-based interactions, we asked 30 breast cancer survivors to share their Facebook pages with our research team. Participants completed a short intake evaluation that included demographic information (sex, racial identity, language, family situation, and income), date of cancer diagnosis, date of transition off of cancer therapy, self-reported level of Facebook use (light, moderate, or heavy), and any observed changes in self-reported Facebook use following their cancer diagnosis or their transition off cancer therapy. In accordance with institutional review board approval, we asked participants to send a friend request to our research account to provide access to profile and page content.

Inclusion and Exclusion Criteria

To be included in the study, participants had to have been diagnosed with stage 1, 2, or 3 breast cancer and out of treatment for at least three months. Participants had to have an active Facebook account for at least three months before diagnosis, though we did not specify a minimum threshold use. The Facebook profile needed to be set up three months before their initial date of diagnosis because we were interested in both diagnosis and transition off cancer therapy. We included only women who had been diagnosed with stage 1, 2, or 3 breast cancer and received chemotherapy treatment. To study both transitions, we did not include patients who had experienced a

cancer relapse. Despite meeting all inclusion criteria, patients were ineligible for the study if they failed to complete the initial intake evaluation or were unable or unwilling to send our team a *friend request* on Facebook—as failure to do so prevented us from accessing the information on their Facebook walls.

Participants

Our final sample comprised 30 women who had been diagnosed with breast cancer between 2010 and 2017. Our sample was exclusively white and female, and all 30 participants reported English as their primary language. The women in our study were aged between 32 and 63 years, with a mean age of 47 years. Overall, the majority of our participants (23/30) reported living with a partner or spouse. Of our 30 breast cancer survivors, 24 reported having at least one child. Study participants generally reported some college education (27/30). With regard to income, whereas 3 women refrained from reporting income, the remaining 27 participants were relatively evenly distributed between the 8 income categories listed in our intake evaluation. Median household income was between US $80,000 and US $99,999, and modal scores were split evenly between the US $60,000 to US $79,999 and ≥US $140,000 income categories. Demographic variables can be found in Table 1.

Table 1. Demographics.

Measure	Frequencies (n)
Age (years)	
30-39	8
40-49	10
50-59	7
≥60	5
Partnership status	
Partnered	23
Not partnered	7
Number of children	
0	6
1	4
2	13
3	5
≥4	2
Income (US $)	
Below 20,000	2
20,000-39,999	2
40,000-59,999	4
60,000-79,999	5
80,000-99,999	4
100,000-119,999	3
120,000-139,999	2
≥140,000	5
Prefer not to say	3

Procedure and Coding

Capitalizing on Facebook's unique timeline feature, we scrolled back to participants' diagnosis date on their Facebook walls. To gather the Facebook data, we used a Google Chrome browser extension called Scraper that allows for specific sections of a viewed Web page to be copied directly into designated cells of a Google spreadsheet. The process involved navigating to the starting date of a given participant's timeline, scrolling and allowing posts to load, and then activating the scraper extension to copy the loaded posts into a spreadsheet. We downloaded a total of 60 files: 2 for each of our 30 participants. One file included data for the 6 months centered on participants' reported date of diagnosis, whereas the other for the 6 months surrounding participants' transition off cancer therapy. Variables automatically downloaded included the following: participants' and poster's name; time, date, type (status update, picture, meme, or video), and the text of the post; and network repose characteristics such as number of likes, shares, and comments the post received.

Analysis

In addition to the variables downloaded automatically by the Web-based scraping program, we were interested in quantifying support exchanges as a proportion of the total number of exchanges on participants' Facebook walls. To evaluate that, our team conducted a line-by-line coding of all status updates and wall posts collected in our data scrape. Our coding team met initially to discuss social support definitions and how social support may be exchanged in social media environments. The coding team also shared academic articles and agreed on the original variables. Our early, theory-based coding scheme was coarse, including dichotomous variables for support exchanged and post valence, along with a categorical variable for type of support exchange.

In meeting 1, each member of the team was assigned a diagnosis and a termination file to review and code. We then met the subsequent week to discuss challenges in data coding and review any status updates or wall posts that could not be neatly categorized based on our 3-variable coding scheme. As challenges emerged with particular status updates, or more generally in observed trends and patterns, we revised our coding scheme and modified and recoded all previously coded data. We repeated this process weekly. Following the coding of our first 10 participants, we had settled on our final coding scheme that was used for the remaining 20 participants. However, our team continued to meet weekly to discuss any difficult-to-code or attribute status updates or wall posts. These were discussed with reference to our original research questions, and in all cases, we were able to reach a consensus.

Once all 30 participants' data had been coded, each team member recoded 1000 lines of the data. These recoded status updates were then compared with the original codes to ensure high inter-rater reliability scores. Given the close collaboration of our team, our inter-rater reliability met all threshold requirement, with percent agreement between 75% and 95% across all variables and Maxwell random error scores between 0.74 and 0.93.

Results

Overview

Our first theoretically grounded coding scheme included all variables automatically downloaded by our browser extension, along with dichotomous variables for valence and support exchange and a categorical variable for type of support. Engagement with the data revealed a number of features of social media data that did not lend to neat theoretically grounded categorizations. The result of this engagement with the data was an iterative adaptation of established definitions of support to capture the variability inherent in Facebook data. In this section, we present our initial variables, challenges that arose from engagement with the data, and resulting modifications to our coding scheme. We present our variables in the order in which they appear in our database: post originator, post content, post valence, support provided, support requested, type of support, and response metrics, followed by a graphical depiction of our coding scheme and a presentation of the final database.

Poster

Although approximately 79.34% (16,912/21,291) of top-level posts were written, posted, or shared by the participant, the remaining 20.56% (4379/21,291) were generated by Facebook friends as either posts made directly to the participant's wall or updates, photos, or videos in which the participant was tagged. Participant-initiated posts were tallied separately from friend-initiated posts, and friend-initiated posts were used to assess the frequency of unsolicited social support interactions. Non–patient-initiated posts were not included in tallies of network response, which our team tallied using quantifiable measures of network response including the number of likes, comments, and unique commenters.

To evaluate changes in support-seeking behavior and network response, we restricted support transactions to those in which the patient was either receiving or requesting support. To contain support exchanges to those focused on the patient, all posts written by the patient in which support was transacted were coded as *support requests*. Posts written by a Facebook friend where support was transacted were coded as *support provision*. Tracking posts as patient- or friend-generated enabled us to exclude posts in which friends asked patients for support or where the patient provided advice, information, or resources to another member of their Facebook network. In addition, given that posts related to a patient's own cancer could come from both a patient or a friend, using the originator of the post alongside the patient's own cancer variable (discussed below) helped us distinguish self-disclosures (patient-generated or own cancer–related) from other discussions of the patient's cancer (friend-generated or own cancer–related).

Distinguishing between patient- and friend-generated content may appear self-evident, but establishing authorship was not always clear. Patients' Facebook pages often included content, memes, blog posts, and articles from other sources, and that content was often accompanied by an introduction or commentary generated by the patient. In an example, a patient posted a meme with a quote attributed to Kathy Kinney:

> *One day she finally grasped that unexpected things were always going to happen in life. And with that she realized the only control she had was how she chose to handle them. So she made the decision to survive using courage, humor and grace. She was the queen of her own life and the choice was hers.* [Patient 23, aged 37 years, diagnosed in 2015]

The meme was accompanied by a post by the patient relating to her own cancer:

> *This gave me strength as I started by new treatment today! It went so well that I left feeling way better than I'd started. Though that was likely due to the awesome [friend] and [friend] and their abundance of snacks…* [Patient 23, aged 37 years, diagnosed in 2015]

In this case, the content of the post originated in part from a second-party source, but the patient was responsible for the introductory text and stood behind the words as an expression of her own feelings with regard to coping and found relevance

to her own experience with breast cancer and treatment. Another illustrative example comes from the husband of one of the breast cancer patients in our group. He logs onto the patient's Facebook page to update the patient's Facebook networks regarding her health status following surgery:

> *Staying overnight following [patient's] surgery. She is doing well but as expected is in quite a bit of pain. Your prayers will surely help her through the night. "Like" this post to let her know that you are with us.* [Patient 19, aged 48 years, diagnosed in 2013]

These instances typify some of the problems faced in attributing content to individual users.

In these specific cases, the first example would have been attributed to the patient, whereas the second instance would have been eliminated from the analysis. These decisions were made with two research goals in mind. The first is that if the patient provides any text or advocates any sentiment be shared or expressed on her Facebook page, they can be seen as attributable to the patient. The second goal was to under-, rather than over-report, support exchanges through Facebook. Specifically, although the husband in the second example was disclosing information, and perhaps issuing a tacit request for support following surgery, the patient was not a part of the support exchange dyad, and thus, although there might be some residual benefit from the likes, responses, and comments, they were not in response to a direct request from the patient. In addition, through a different lens, the post from the husband could be seen as providing support to fellow friends and family who were concerned about the patient's health status following surgery.

Problem-Related, Cancer-Related, or Own Cancer–Related Post Content

Organizing cancer-related posts resulted in another important distinction within the category of cancer-related posts. We surmised early on that it was important to establish whether posts related to a problem or not. However, problem-related posts were not systematically cancer-related, and posts that were cancer-related were not systematically related to the patients' own cancer. The goal of including a general *problem-related* variable rather than focusing exclusively on cancer-related posts reflected an objective from the outset to assess whether those who posted more often about problems before cancer diagnosis were more likely to post about cancer-related problems following diagnosis. In addition, patients often posted about non–cancer-related problems both before, during, and following the treatment. These posts discussed a gamut of issues both health-related and otherwise. For example, in the three months following breast cancer diagnosis, a patient posted:

> *Ugh. I got an IUD today and ladies who know, the pain suuuuucks...but I feel better knowing I'm ok to continue treatment with peace of mind. And I'm glad I didn't have to pay for that! Ouchie :-(* [Patient 36, aged 35 years, diagnosed in 2013]

Another patient posted about job benefits, whereas a third posted a status update about the health problems her cat was experiencing:

> *Anybody experienced with kidney disease in cats? Our little Twinkie is not well. Poor sweetie. She's only 4. This. Sucks.* [Patient 14, aged 45 years, diagnosed in 2013]

Distinguishing between problem-related, cancer-related, and own cancer–related posts will ultimately enable us to explore whether patients who post more about their problems before cancer diagnosis also post more about their problems following cancer diagnosis or other health-related trauma. In addition, by highlighting problem-related posts that did not related to cancer, we set the stage for an analysis of differential support availability following diagnosis. Specifically, do posts about cancer-related problems get more attention from Facebook friends than non–cancer-related problems?

An added complication in coding posts as either cancer- or non–cancer-related was that frequently *cancer-related* posts related to something other than the patients' own cancer. Following cancer diagnosis, a number of participants became active in breast cancer campaigns—raising funds for breast cancer research and collecting supplies to send to such patients in the hospital. Breast cancer survivors in our study participated in both walks and races with fellow breast cancer survivors and made connections that ultimately impacted their Facebook network composition. As a result, several of the participants' pages included pictures of fundraising activities, articles, or opportunities for friends and family to get involved either as potential donors or participants. Although these posts were cancer-related, and fundraising efforts were catalyzed by patient's own diagnosis, the posts had very little to do with the patients' own experience with cancer. We argue that the degree to which a post related to a patient's own cancer experience would likely have implications for resulting support provided and, therefore, opted to include dichotomous variables for both *cancer-related* and *patient's own cancer*.

All posts that related to the patient's own cancer were coded as problem-related, cancer-related, and own cancer–related. This categorization allowed us to keep a register of the concerns and issues raised by the patients with respect to their own cancer, even when no support was requested. We labeled these posts as *self-disclosure* and interpreted them as tacit requests for support, in accordance with the study by Zhang et al [51]. This register of self-disclosures enabled us to index the types of issues discussed by the cancer patients. Self-disclosures could be general or specific and handled a range of issues from personal challenges to triumphs and updates on disease progression and cancer treatment. For example, a participant posted:

> *All 3 cousins got cancer before age 50. Unfortunately for two cousins their cancer was not detected early. Next week when I ride the Canary Challenge I will be celebrating my own "getting through cancer" as well as riding in memory of my special cousins...whose cancer was too far gone—sadly [cousin 1] passed at age 48 years, [cousin 2] died just a few months short of 53 years. I miss them both a lot.* [Patient 24, aged 52 years, diagnosed in 2013]

The post is quite general and focuses outward on others' cancer experiences but, nevertheless, includes a brief mention of the participant's own cancer experience as the motivator for participation in the Canary Challenge. Other self-disclosures were more specifically focused on the patient's own experience with cancer diagnosis and treatment:

> *Hi all! I just went for my first walk around the ward. It went very well! Everyone here seems very happy and even impressed with my progress so far. I'll be headed home this afternoon! ...thank you all for your overwhelming love, support, and encouragement!!! The docs say my lymph nodes are clear, so this chica is CANCER FREE!* [Patient 26, aged 35 years, diagnosed in 2013]

The second status update appears to serve a dual purpose of (1) updating friends and family on Facebook with regard to treatment progress and (2) inviting them to celebrate a treatment milestone, though the patient never explicitly requests support.

Valence

We characterized all posts related to a patient's own cancer as *problem-related* under the assumption that there was no positive development related to cancer (symptom abatement, treatment, or progression) that was not overshadowed by the problem of having the disease. In other words, rather than *good* and *bad* developments in cancer, the range could more reasonably be considered as varying between more and less bad. Evaluating valence, on the other hand, allowed us to distinguish between positive and negative developments in cancer and its progression. Patients could talk about developments in their cancer symptomology or treatment but put a more or less positive spin on them. For example, a patient provided updates on her progression through cancer treatment by including a countdown:

> *Health update! Radiation begins tonight. Only 7 more weeks to go!* [Patient 23, aged 37 years, diagnosed in 2015]

Another patient updated:

> *I heard from my doctor this week regarding the cancer gene- I do not have it. All tests came back normal so we can praise God for some good news. She did say because of the amount and location of lymph nodes involved they are considering me stage 3 so we are very thankful my treatments have started so soon and they are being so aggressive. Even though the nausea is not letting up, I'm remaining positive that the cancer is being killed. Thanks for the continued prayers and support.* [Patient 38, aged 33 years, diagnosed in 2016]

Whereas the valence in the first post is largely positive, that in the second post is more mixed. It includes both positive and negative developments in disease and treatment progression.

To decide on how to handle mixed-valence and other complicated posts, we went back to the research questions underpinning the inclusion of the variable in the first place. In this case, we opted to include a category to assess valence for two reasons. The first is that change in valence has been used to assess the quality of support exchanged in Web-based communities [52]. As a result, it is useful baseline information to have for any study that discusses the nature, quality, and particularly the effectiveness of online social support coded from Web-based social interactions. The second is that research has demonstrated a number of caveats to the positive relationship between online support exchange and improved health; and among other things, patients have cited forced positivity as the potential to diminish the quality, and by extension the health buffering effects, of supportive exchange. Thus, we were interested in whether Web-based social media systems, such as Facebook, were likely to reward patients whose cancer-related posts were characterized by more positive sentiment.

Despite a clear definition of valence, assessment was not straightforward. Some posts were characterized by clear valence. For example, a patient posted a *rant* about the unsolicited advice she received regarding self-care and cancer treatment:

> *Day 4 nicotine free! Saw my Mom's fam tonight and told those who didn't know about my current situation. It went well, no tears from me, and best of all no really upsetting "well, should you really do that…" TRUST me, this is moving fast for me too BUT it's also my life, my choice. Please leave the second guessing to ME!!!! Rant: over.* [Patient 36, aged 35 years, diagnosed in 2013]

Posts of this nature, conveying a very clear valence, were comparatively rare. More typically, posts included a gamut of positive or negative emotions. Sometimes the positive and negative emotions were mixed together, as in the above quote. In other instances, posts began with a more negative or neutral disease progression update and ended with an expression of determination or gratitude. For example, a participant posted:

> *I am currently 1 week out from my second chemo treatment. I am 27 years old and have a 2 and 1 year old. You never know when it will be you. My mother always taught me to check and I found my triple negative by myself. Bless all those who have faught [sic] or who are fighting right along with me.* [Patient 26, aged 32 years, diagnosed in 2012]

The post begins with objective or possibly neutral information but ends with an expression of gratitude, first to the patient's mother and then to other cancer patients. Guided by our notion of *forced positivity* in face-to-face interactions [36], we opted to interpret the inclusion of positive sentiment or gratitude as possible evidence of that patients felt compelled to shift the valence in their posts to receive a response from friends and family through Facebook. As such, status updates characterized by mixed valence were coded with respect to the final sentiment expressed in the post.

Support Provided

Posts could only be coded as providing support if they were generated by a Facebook friend. Given that not all support exchanges that transpired between cancer patients and their support networks were support provided to the patient or support requests from patients, it was important to establish which

support exchanges were most meaningful in the context of changes in support-seeking behaviors following breast cancer diagnosis. As such, we excluded any instances in which breast cancer patients were asked to provide support. In addition, coding top-level posts for evidence of support provision enabled us to establish the frequency of unsolicited support and measure it by support type: informational, resource, emotional, or general advice.

In our sample of 21,291 top-level posts coded, 5178 posts were unsolicited support posts from friends. As shown in Table 2, posts were roughly evenly divided between those around the time of cancer diagnosis (N=2457) and those surrounding transition off cancer therapies (N=2721), but there was a considerable 3-fold increase in unsolicited support provision immediately following cancer diagnosis.

Table 2. Code count and percentages of support requests.

Support request and types	Patient			Friend		
	Diagnosis (N=9322), n (%)	Termination (N=11,223), n (%)	Total (N=20,545), n (%)	Diagnosis (N=2457), n (%)	Termination (N=2721), n (%)	Total (N=5178), n (%)
Support requests	341 (3.66)	291 (2.59)	632 (3.08)	748 (30.44)	725 (26.64)	1474 (28.47)
Support type						
Informational	53 (0.57)	36 (0.32)	89 (0.43)	22 (0.90)	27 (0.99)	49 (0.95)
General advice	18 (0.19)	19 (0.17)	37 (0.18)	1 (0.04)	1 (0.00)	2 (0.04)
Emotional	88 (0.94)	71 (0.63)	159 (0.77)	607 (24.70)	585 (21.50)	1192 (23.02)
Resource	201 (2.16)	172 (1.53)	373 (1.82)	144 (5.86)	145 (5.33)	289 (5.58)

It is worth noting that among friend-generated wall posts, not all were intended to provide support. In addition, instances of support provision only accounted for 698 of the total 1329 total increase in posts between the three months preceding cancer diagnosis and the three months following it. This supports the notion that friends were generally checking in more with cancer patients following cancer diagnosis than they had been before diagnosis, though not always around cancer or expressions of social support.

Observations similar to the one above complicated the coding process with respect to support provision. To simplify, we operationalized support as exclusively those posts that contained encouragement, esteem support, information, resources, or the intention to transmit resources (for a full description of the definition of resource support, see the *support types* section below). Nevertheless, support provision was complicated in that it required coders to pay close attention to both macro- and microlevel contextual cues that could indicate whether a post was meant to provide emotional support and support was specifically cancer-related. In other words, in some instances, patients' family members would post from their own account to provide information and rally support for the patient. The following example is from one of our participant's daughters:

> *Tomorrow my mom will be having her surgery. Please continue to pray for her. These past six months have been difficult for her and our family and I appreciate all of the support everyone has given us. She has put up quite the fight and continues to keep strong. I love you, mom!* [Daughter, patient 30, aged 63 years, diagnosed in 2013]

This post provides encouragement and information on behalf of the patient and so despite not containing a direct message of support, it would have been coded as support provided. Conversely, friends often posted to participants' walls general words of encouragement and support, such as *you rock!* Owing to the absence of indicators that the post was specifically meant to provide encouragement in the fight against cancer, we opted to estimate conservatively, calling the posts non–cancer-related transmissions of general emotional support.

Our sample also included several lengthy posts expressing support for cancer patients made to the friend's own wall and simply tagging the cancer patient. On Facebook, when an individual creates a status update, they are able to tag a second Facebook user if that user is a friend. In so doing, the post will appear on both the original poster's wall and that of any person tagged in the post. Our data showed a number of instances where a friend wrote a post about cancer and tagged a group of people that included the patient. One such instance came from a fellow breast cancer patient who tagged a group of 88 friends and posted:

> *Hopefully I tagged everyone! It's been a hell of a year so far but I am very grateful for you all! I'm feeling so much better and ready for Chrissymas!!!* [Friend, patient 25, aged 51 years, diagnosed in 2014]

In these instances, it was unclear whether to interpret the post as providing support in that the original post was not directed at the patient herself. However, we opted to code the posts as both related to the patient's own cancer and transmissions of actual support in that these were designed to express solidarity. It may have been an unwelcome expression of support or broadcasting of a particular patient's health status, but the design of the post was to express support that was not directly in response to a request from the patient.

Microlevel context was also important but often quite difficult to assess. For example, a patient posted frequent chemotherapy countdowns. As a result, many of her network were aware of when her final chemotherapy was to occur and stopped by her page to offer words of encouragement. The words of encouragement never mention cancer specifically. For example, a participant posted a countdown to her final chemo treatment. As a result, on the day of her last treatment, she received

numerous posts congratulating her—some referencing her final treatment, and some did not. One user wrote a note of congratulations and support but never mentioned cancer directly:

> *What a long hard ass journey and trail this has been. Not over yet, but the big hurdles are not behind us! We are blessed that we found the C when we did, for if not...things would not be the same right now. I have learned so much about myself, and strive to be the best husband I could be... Love you honey and congratulations!* [Patient 16, aged 56 years, diagnosed in 2013]

As in the above example, the nature and content of the post as well as surrounding posts made it clear that these were made in celebration of an impending completion of cancer treatment. In addition, in a number of instances, the support provided directly to a patient's page by a friend was in response to a direct request for support from the patient. This type of support exchange occurring across top-level posts was quite rare. In addition, separating solicited and unsolicited support in top-level posts was beyond the scope of this study. As a result, we interpreted all friend-generated support-providing posts as examples of unsolicited support exchange.

Support Requested

A common feature of research around Web- and non–Web-based support exchange is that support needs can be expressed both directly and indirectly. Indirectly, support needs may be expressed through the disclosure of personal information. Directly, individuals may ask for advice, information, or resources. As we had already parsed out problem-related posts from general cancer-related posts and posts related to a patient's own cancer, our assessment was that we had captured the tacit support requests. As a result, we only coded status updates as *support requests* if the status update went beyond disclosing a need and actually issued a direct request for advice, encouragement, information, or resources. This greatly restricted the number of posts coded as support requests, but the decision was made to separate direct and indirect request in that such a separation enabled us to view differences in support availability on the basis of whether the post included a direct request for support or relied on more indirect methods of communicating a support need.

An additional complication that emerged in the coding of support requests was the requesting of support on behalf of a cancer patient. Posts were often posted directly to the friend's wall but appeared on the patient's wall because she had been tagged. Tagging the patient offers the benefit of ensuring wider broadcast; when a person is tagged in a post, the post is broadcast to not only the poster's social network but also that of the patient. Although it was clear that the support requested was intended to benefit the breast cancer patient, it was unclear from these second-party support requests whether patients ever asked their friends to issue these calls for support. These posts differed from the general cancer condemnations in the *Support Provided* section above in that the posts made an express request for support on behalf of the cancer patient, rather than a simple condemnation of cancer. In the context of second-party posts, our team made the decision to count them as a provision of resource support in that they were allowing cancer patients to access support from a broader network of support.

Type of Support

Typology Overview

If a post included a direct request for support or evidence of support being provided, we categorized the post by the type of support using established categories from social psychology literature [23]. Cancer patients are known to follow certain patterns in their support-seeking. Specifically, they tend to be selective about the source of various types of support—seeking emotional support from friends and family but preferring that informational support come from doctors or other breast cancer patients and survivors. Online support-seeking tends to follow a different pattern. Likely given the ease of transmission through the internet, research on computer-mediated communication suggests that, broadly speaking, users tend to gravitate toward emotional and informational support when exchanging support on the Web [36,53]. Beyond advancing research on the nature and quality of online support exchange, knowing the frequency with which various types of support exchanged could also help advance our understanding of cancer patients' online support-seeking behavior. Specifically, do breast cancer patients behave more similar to other cancer patients in their support-seeking, or they seek the same types of support as other internet users?

To assess this, we began with a typology of three types of support as outlined in the theoretical framing of the paper. Posts that were assessed as either providing or requesting support were then evaluated for the type of support that was exchanged in the post. Following broad definitions laid out in the review of relevant literature above, we assessed whether each support exchange included transfer of information, resources, or socioemotional support. The categories are defined, and examples are provided in Table 3.

Table 3. Categories of provided and requested support, including definitions and examples from the data.

Support Type	Definition	Provided	Requested
I (Informational)	Transfer of relevant information to help cope with a problem	Hey lady…happy "last chemo" day! Came across a website…about all sorts of freebies, from wigs to housecleaning. Thought I would pass it along. (Patient 36, age 35, diagnosed 2012)	Calling all cooks – what's your favorite Vegetarian/Vegan meals? I know there's some of you out there. I am hoping to make more vegetarian meals in 2012. Stuff that's kids like is a double plus! (Patient 13, age 41, diagnosed in 2011)
A (Advice)	General advice, not related to factual information	No instances of unsolicited general advice provided	Has anyone seen the new jungle book movie in the theater? Wondering if it's too scary for [child]. He's never been to a movie theater and I thought this movie might be a good one. (Patient 38, age 33, diagnosed in 2016)
E (Emotional)	Feeling of togetherness or the knowledge that one is valued	Miss [patient's name]! How are you darling…it's been a looooooonngggg time! Your little man is so adorable. Not sure what's going on, but wanted you to know that I'm praying for you. Grab your strength from God…He provides us with all we need. Thanks for adding me! (friend, Patient 26, age 32, diagnosed in 2012)	Labs came back really low today, so if you've been sick or by someone sick please stay away. Say a prayer they go back up before my next treatment and that I get some energy. I've been exhausted the past two days, probably too much fun. (Patient 36, age 33, diagnosed in 2016)
R (Resource)	Actions and materials made available through individual support network	Consider helping this beautiful young mama, [tagged patient], if you can. Every little bit helps. Read her story. BREAST CANCER/ALL CANCER SUCKS!! (friend, Patient 29, age 35, diagnosed in 2016)	[Daughter]'s first day of school is tomorrow, and she's very nervous…she's coming in during the middle of the year. If you are her friend on fb or have her Snapchat…go flood her with encouraging messages please. Everyone else please say prayers for confidence, comfort, and nice kids to become good friends. Thanks! (Patient 29, age 36, diagnosed in 2016)

Emotional Support

Early on, our team made the decision to consolidate esteem support and social companionship under the auspices of emotional support. This is consistent with other coding schemes, as the distinction between esteem support (that one is loved or esteemed) and social companionship (that one is not alone) adds less in this context than a broad-level emotional buffering. Emotional support constituted the largest number of unsolicited support posts made by friends. In cases of support provision, the posts were very succinct and included statements, such as *you rock* or *you are so strong!* Longer message may be more specific to the person but include similarly general sentiment. For example, a friend of patient 17 posted the following to her wall:

> *Hey girilie [sic] Just wanted to stop by and say hi. I hope recovery is going great and hopefully alot [sic] less painful today then [sic] yesterday!!! Been praying for you:) How have you been otherwise, I see that you are blessed with two little ones they look so much like you:) Congrats on being a mama even though it's a couple years late lol. Well hit me up whenever, I'm sure you'll be online a bit more offten [sic] while you recover and I work in frunt [sic] of a computer so I am always on:) Have a great rest of you [sic] day and just know I'm praying for everything...fast recovery, little less pain eachday [sic], and yummy food made buy [sic] your honey for you:) Love yea Kim.* [Friend, patient 17, aged 32 years, diagnosed in 2012]

Despite the longer post content, the sentiment is quite similar to the shorter *you rock* type posts, offering general encouragement, a sense of togetherness, and small talk but limited engagement with the patient or their cancer.

In the context of support requests, emotional support was often requests for *thoughts*, *prayers*, *thoughts and prayers*, and even *positive vibes*, *mojo*, *karma*, or *luck*. In most instances, when the requests related to patient's own cancer treatment, they were attached to informational updates related to cancer treatment or specific challenges. For example, a patient provided an update on her health status and specifically implored that friends kept praying on her behalf:

> *Thank you all for the positive thoughts and prayers today...I felt them, even if I didn't get the results I wanted. Doctors have some additional testing they want to do so they can determine the type of chemo I will receive before surgery. So...more waiting but hopefully tests will come back fine and we can proceed. Keep those prayers coming please.* [Patient 22, aged 39 years, diagnosed in 2015]

In addition to requests for thoughts and prayers around their own cancer treatment and recovery, cancer patients also requested emotional support for both general cancer– and non–cancer-related issues, including emotional support requests on behalf of others. For example, one of our participants requested emotional support for a friend:

> *[Friend] just left for his angiogram which he didn't want to tell anyone about, but I can tell people so I can ask you to please keep him in your thoughts?* [atient 14, aged 45 years, diagnosed in 2013]

It is important to note that for the women in our study, emotional support exchanges extended beyond both themselves and their cancer to include others in both their Web- and non–Web-based friend networks.

Informational Support

Informational support exchange through Facebook was exceptionally limited with only 159 requests for information and 49 instances of unsolicited support–providing posts. In addition, almost none of the 208 informational support exchanges in our 21,291 total coded posts included information about cancer, its treatment, side effects, lifestyle, or diet. In this case, cancer patients on the Web tend to behave much more similar to those outside it in their preference for emotional and resource support and reluctance to exchange information through Facebook.

Resource Support

There were two primary types of resource support requests that featured on cancer patients' Facebook walls. The first were awareness or fundraising campaigns. Though we did not distinguish between the two categories in our coding, this type of awareness and fundraising campaign post might best be termed broadcasting posts. Many of the patients in our sample became very active in raising funds and awareness for both cancer in general, pediatric cancers, or breast cancer subsequent to their cancer diagnosis. Patients would then use their walls as a space to broadcast information on fundraising initiatives to friends and family. For example, a patient campaigned consistently to raise money for childhood cancer and used her Facebook page to raise both awareness and funding for the foundation:

> *Guess what? If 17 of y'all gave up drinking one green beer or cup of coffee this weekend ($5 only) I would reach $3300 - and be that much closer to helping find a cure for kids cancer. This is your last chance to help support me before I shave my head tomorrow!* [Patient 13, aged 41 years, diagnosed in 2011]

In addition, several cancer patients sought to raise cancer awareness or share stories of other cancer patients who were either in treatment for cancer or had died following cancer treatment. To that end, many patients used their Facebook pages to request that their friends *spend a minute* visiting a Web page, clicking a link, or reading the story of a fellow cancer patient. As these requests were not always information based, we opted to conceptualize time as a resource and code these requests as resource support requests.

A second category of resource requests corresponded more closely with two common conceptualizations of resource support in social psychology literature. In these cases, patients made direct requests for the resources that they needed to attend to obligations both in and out of the hospital. Some of the materials were things that the patients needed to be more comfortable while in the hospital, such as the patient who requested socks. Other requests centered on resources that could be transmitted over the internet. For example, unable to attend her child's recital, a patient rallied support through Facebook to obtain photos or videos of the event:

> *Since I am missing the concert this morning, if anyone takes video's [sic] or pictures and captures my kiddos could you send them my way, please.* [Patient 9, aged 45 years, diagnosed in 2014]

These requests were more specific to a patient's own situation and directed at particular members of her friend network who were local and attending the school concert. As a result, it may make sense for future iterations of the coding schema to distinguish them from general fundraising, awareness raising, and broadcasting posts.

Resource support providing posts were similarly divided. Broadcasting posts were those in which friends reposted a cancer patient's request for fundraising or other types of support. As mentioned before, this type of support was seen as allowing the cancer patient to access a new network of individuals with requests for support and required time on the part of the support provider. A patient's fundraiser in support of children with cancer was spread by several of her friends:

> *We are heading downtown tomorrow night to support our friend Naomi Bleecker Damask while she shaves her head to raise money for kids battling cancer through St. Baldrick's. She is doing this for other kids while in the thick of her own cancer battle, so couldn't we all take 5 minutes out of our busy day to donate just $5 or whatever you can spare to help a child who is battling cancer. Thank you!!!!!!!!!!* [Friend, patient 13, aged 41 years, diagnosed in 2011]

In addition, there were several instances in which a friend, often a partner, provided an update on the cancer patient's health status. Although naturally, this would fall under informational support between the patient's partner and her friends, this was coded as another resource support provision under the broadcasting category in that it required time on the part of the provider and was a service (whether solicited or unsolicited) provided to the patient.

Another complication that arose from the data was how to manage the fact that most of the resources requested by a patient or provided to the patient could not be transmitted over the internet. When the friend of a cancer patient offered to pick the patient's child up from school and drop the child off at home, this was a clear intention to provide resource support but did not include the transmission of any actual resources. For purposes of this study, we coded any intention to share resource support as a resource support exchange. Proportionally, it is likely that not all of the intentions expressed actually manifested in the transmission of resources, but the intention to transmit resources gives researchers an idea of who is willing and able to provide those resources, and it might be interesting to evaluate the degree to which intention to transmit resources translates into actual transmission of those resources.

General Advice

Some posts made by patients were clear requests for engagement but did not include a request for emotional, informational, or resource support. Although infrequent, these posts offered a situation or choice and asked Facebook friends to offer advice on the choice to be made. In an example, a patient asked:

> *Another house question: double sinks with little counter space or 1 sink with more counter space? I'm wanting the counter space. I only need the sink for two minutes.* [Patient 38, aged 33 years, diagnosed in 2016]

These were often not cancer-related nor did they pertain to any factual information—rather, solicited the opinions of friends and family connected through Facebook. As there was no preexisting typology available in the support literature to categorize the request, our team opted to create a fourth, *general advice* category. This category subsumed any general solicitations of the opinion of the Facebook network in general.

Response Metrics

In this phase of the project, we opted to hand code only the 21,291 top-level posts. The approximately 80,000 remaining posts could provide unique insights into the nature and quality of network response to both support requests and provision, and coding schemes for these remaining responses are currently underway. In the meantime, metrics such as the number of likes and comments have been used to assess network response to support requests. However, those metrics were not available, as Facebook only had a *like* button until 2018. This created complications for acknowledging a significant event, observation, or disclosure if the post was characterized by negative valence. In addition, the number of responses had the potential to be complicated by instances in which friends began talking or arguing in the comments section of a specific post. The result of the latter could be a post with multiple comments—possibly interpreted as significant and broad support exchange but included no exchange with the actual person posting. To remedy this, we created an algorithm to count the number of unique commenters. In this way, we could look at the number of comments with respect to that of unique commenters to assess whether a conversation had developed, or the support came from a number of different sources. It is noteworthy that few sources of many comments may not indicate low-quality support if the patient starts a more involved conversation with a friend in the comments section of a post; it may simply indicate that the nature of the support exchanged is likely to be different, and such differences warrant mention and additional attention in analysis.

Timestamp

In addition to understanding the changes in each patient's support-seeking behavior following cancer diagnosis, we wanted to look for patterns that were consistent between patients and assess the degree to which support fluctuated over time and with respect to other transitions in health status. To do this, our team used a Unix timestamp for each post and comment. We also used a Unix timestamp for patient's date of diagnosis and their transition off cancer therapy. This allowed us to superimpose all 30 patients' 6-month Facebook posting trajectory to view broad-level trends in our sample's posting behaviors across time to highlight any patterns in fluctuations that might serve as rich points for additional qualitative and quantitative analysis.

Data

Figure 1 graphically illustrates how our coding scheme was developed and refined based on observations from the originally downloaded dataset, and Figure 2 shows our final coding scheme. Tables 2 and 4 provide a broad overview of the resulting dataset and include disaggregation by (1) patient, friend, and total posts, (2) proportion of total posts initiated by the patient versus her Facebook friends, (3) post content, including problem-focused, cancer-focused, and patient's own cancer–focused posts, and finally, (4) proportion of posts that included support exchange broken down by information, advice, resources, and emotional support.

Figure 1. Final coding scheme progression.

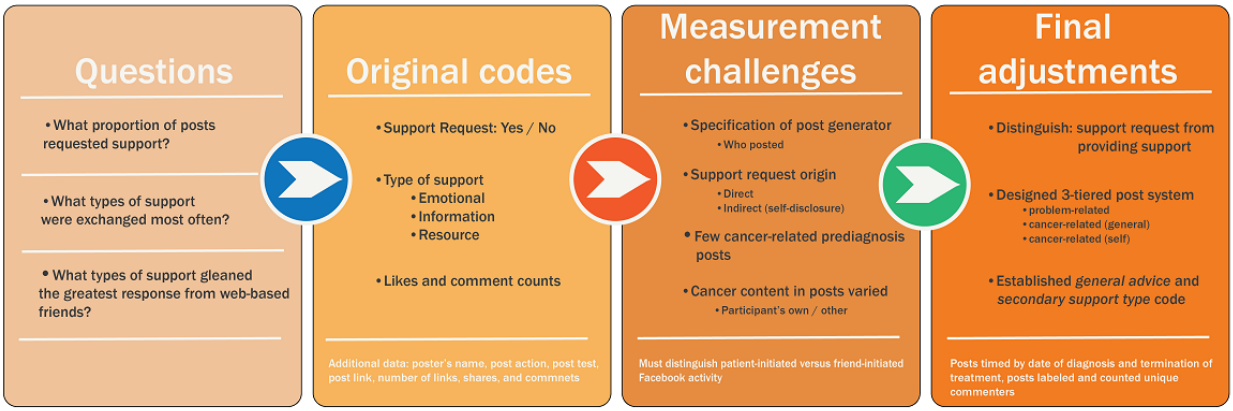


Figure 2. Final coding scheme. I: informational; A: general advice; E: emotional; R: resource.

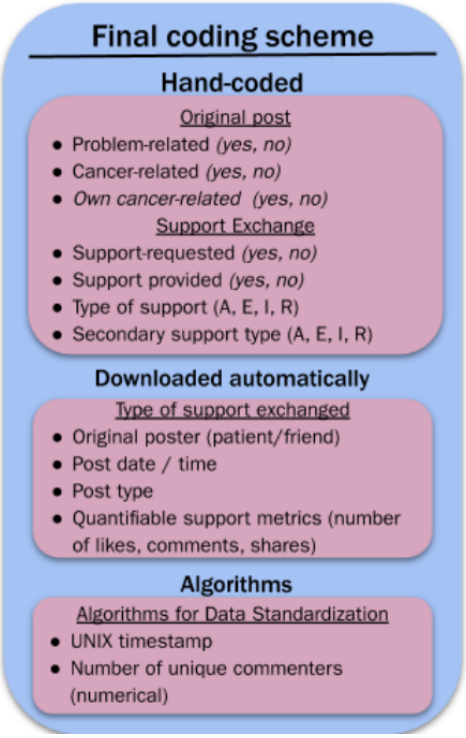


Table 4. Summary statistics—posts.

Categories	Friend posts (N=4379), n (%)	Patient posts (N=16,912), n (%)	Total posts (N=21,291), n (%)
Problem-related	664 (15.23)	2006 (11.89)	2670 (12.57)
Cancer-related	920 (21.09)	1889 (11.19)	2809 (13.22)
Patient's cancer	720 (16.51)	1358 (8.05)	2078 (9.78)

From the tables, we see limited requests for support and a mismatch between the types of support being requested and that being provided by friends. Although patient's own posts account for the large majority of overall posts, the overall proportion of patients' posts that include a request for support is only around 3%, with self-disclosures at around 8%. In addition, over half of patient requests are for resource support (59%), whereas only 27% request emotional support. By contrast, friends' posts were twice as likely to bring up the patient's own cancer and over 9 times as likely to include support. However, despite patients requesting a preponderance of support requests, the overwhelming majority of support provided (80%) was emotional support, with only around 19% offering resource support.

Discussion

Principal Findings

We set out to measure support exchange around two critical transitions in cancer treatment: diagnosis and transition off cancer therapy. Our preliminary review of the more than 100,000 lines of data collected from 30 breast cancer patients' Facebook activity surrounding each transition revealed more questions than answers. Were we interested in support around any and all issues or simply around breast cancer? Will a support request always include a statement of need or want, or can Facebook users request support by simply disclosing information about their situation? How do we count unsolicited posts from friends that contain supportive, if limitedly substantive, statements such as, *you rock*? And perhaps even more muddy, how do we manage situations where participants are tagged in someone else's post or where husbands post updates from their wives' Facebook account? These issues are made all the more difficult by the fact that rationale (and often empirical precedent) exists to justify either decision.

Answering these broader questions began with a survey of the literature and a guiding philosophy that favored theoretically grounded divisions. We began with the data that were scraped automatically from our browser extension and then created variables that would allow us to explore theoretical concepts in support exchange. These included dichotomous variables for support exchange and valence and a categorical variable for type of support exchanged. Finer distinctions emerged iteratively in coding, as our team encountered status updates and wall posts that could not be fit to the original coding scheme. A total of 8 themes emerged in two categories, specifically data comparability and operationalization of support. Comparability refers to grouping similar components. In other words, if we are interested in social support exchanges, it is important to distinguish between posts created by patients and those created by friends, that predate cancer diagnosis, and that do and do not request support. Other distinctions resulted from a difficulty in

operationalizing support and resulted in adapting categories and definitions to suit the nature of the data, from counting requests to read as a request for friends' time (resource support) to creating a new category for general advice to cover posts requesting opinions on hair styles or dresses.

The resulting dataset provides unique insights into the nature and quality of support exchanges via Facebook. Though derived from a smaller sample of breast cancer patients, the database contains 21,291 individual posts from Facebook users and provides unique insights into the utility and responsiveness of social support exchanges via Facebook. Broad-level data categorizations indicate that Facebook may be effective in allowing patients to fulfill social obligations and broadcast general needs but provides social support that is very general and does not respond to any specific or cancer-related support need. To illustrate, despite 16,912 status updates, only 3% of posts included a direct request for support, and the majority of those support requests were for resource support. By contrast, friends' posts accounted for only 20% of overall posts but were twice as likely to include reference to the participant's own cancer and 9 times as likely to include support. However, the overwhelming majority of unsolicited support provided by friends was emotional, with less than one-fifth offering resources—time or services.

Given the importance of social media in social interaction, support exchange, and health promotion, our findings offer a timely contribution to methods of evaluating support exchange in social media environments. Current trends in social media research and Web-based peer exchange show important avenues for future insights on the value of social media to overcome communication barriers normally associated with in-person or more formal and potentially prescriptive modes of social support exchange, but such research has been hampered by a need for consistent operationalizations of online social support that cuts across studies. Where this study offers insight beyond existing comparisons of social support exchange with Web-based economies of information exchange [54], is the way this coding schema permits a more compelling analysis of support exchange as a type of *peer community*, where members act proactively to buffer stress effects associated with negative health experiences.

Reviews of qualitative or mixed-methods approaches to Web-based data analysis point to trends in research using counts and content and thematic analysis; however, few studies show researchers applying strategic approaches to answer specific research questions using Facebook or other social media data [55]. This study addressed two of those issues by providing a systematic coding schema that allowed a characterization of social exchange behaviors to suggest that Facebook use for social support extended beyond information provision and instrumental support to include socioemotional support. In addition to providing clarity on social media behaviors for cancer patients, our paper outlines a clear method and design for organizing and assessing unwieldy social media data in useful ways. Such an approach addresses a gap in the field emphasized by recent review data showing that better research designs and methods are needed for examining the effectiveness of social media platforms for health benefit [3].

Conclusions

Social media data are unwieldy and not always conducive to neat categorizations. Designing a data analysis approach a priori for organizing and systematizing Facebook data allows us to explore the functionality of Facebook as a platform for the exchange of social support. In doing so, this paper provides insights into the different forms that informational, resource, and emotional support may take in Web-based social environments such as Facebook and challenges researchers may face in measuring those constructs and how to respond to such challenges to create broader continuity in evaluation and measurement of computer-mediated support exchange.

Conflicts of Interest

None declared.

References

1. Kaplan AM, Haenlein M. Users of the world, unite! The challenges and opportunities of social media. Bus Horiz 2010 Jan;53(1):59-68. [doi: 10.1016/j.bushor.2009.09.003]
2. Chou WY, Hunt YM, Beckjord EB, Moser RP, Hesse BW. Social media use in the United States: implications for health communication. J Med Internet Res 2009 Nov;11(4):e48 [FREE Full text] [doi: 10.2196/jmir.1249] [Medline: 19945947]
3. Giustini D, Ali SM, Fraser M, Kamel Boulos MN. Effective uses of social media in public health and medicine: a systematic review of systematic reviews. Online J Public Health Inform 2018;10(2):e215 [FREE Full text] [doi: 10.5210/ojphi.v10i2.8270] [Medline: 30349633]
4. Haythornthwaite C. Social networks and internet connectivity effects. Inf Commun Soc 2005 Jun;8(2):125-147. [doi: 10.1080/13691180500146185]
5. Mikal JP, Rice RE, Abeyta A, DeVilbiss J. Transition, stress and computer-mediated social support. Comput Hum Behav 2013 Sep;29(5):A40-A53. [doi: 10.1016/j.chb.2012.12.012]
6. Chee W, Lee Y, Im E, Chee E, Tsai H, Nishigaki M, et al. A culturally tailored internet cancer support group for Asian American breast cancer survivors: a randomized controlled pilot intervention study. J Telemed Telecare 2016 Aug 2;23(6):618-626. [doi: 10.1177/1357633X16658369] [Medline: 27486198]
7. McCaughan E, Parahoo K, Hueter I, Northouse L, Bradbury I. Online support groups for women with breast cancer. Cochrane Database Syst Rev 2017 Mar 10;3:CD011652. [doi: 10.1002/14651858.CD011652.pub2] [Medline: 28278559]

8. Winzelberg AJ, Classen C, Alpers GW, Roberts H, Koopman C, Adams RE, et al. Evaluation of an internet support group for women with primary breast cancer. Cancer 2003 Mar 1;97(5):1164-1173 [FREE Full text] [doi: 10.1002/cncr.11174] [Medline: 12599221]
9. Batenburg A, Das E. Emotional coping differences among breast cancer patients from an online support group: a cross-sectional study. J Med Internet Res 2014 Feb;16(2):e28 [FREE Full text] [doi: 10.2196/jmir.2831] [Medline: 24499687]
10. Schroevers MJ, Helgeson VS, Sanderman R, Ranchor AV. Type of social support matters for prediction of posttraumatic growth among cancer survivors. Psychooncology 2010 Jan;19(1):46-53. [doi: 10.1002/pon.1501] [Medline: 19253269]
11. Drageset S, Lindstrøm TC, Giske T, Underlid K. Women's experiences of social support during the first year following primary breast cancer surgery. Scand J Caring Sci 2016 Jun;30(2):340-348. [doi: 10.1111/scs.12250] [Medline: 26179112]
12. Høybye MT, Johansen C, Tjørnhøj-Thomsen T. Online interaction. Effects of storytelling in an internet breast cancer support group. Psychooncology 2005 Mar;14(3):211-220 [FREE Full text] [doi: 10.1002/pon.837] [Medline: 15386774]
13. Uchino BN. Social support and health: a review of physiological processes potentially underlying links to disease outcomes. J Behav Med 2006 Aug;29(4):377-387 [FREE Full text] [doi: 10.1007/s10865-006-9056-5] [Medline: 16758315]
14. Uchino BN. Social Support and Physical Health: Understanding the Health Consequences of Relationships. New Haven, CT: Yale University Press; 2004.
15. Durkheim E. Suicide: A Study in Sociology. Glencoe, IL: The Free Press; 1951.
16. Berkman LF, Glass T, Brissette I, Seeman TE. From social integration to health: Durkheim in the new millennium. Soc Sci Med 2000 Sep;51(6):843-857. [doi: 10.1016/S0277-9536(00)00065-4] [Medline: 10972429]
17. Dean A, Lin N. The stress-buffering role of social support. Problems and prospects for systematic investigation. J Nerv Ment Dis 1977 Dec;165(6):403-417 [FREE Full text] [doi: 10.1097/00005053-197712000-00006] [Medline: 591939]
18. Granovetter MS. The strength of weak ties. Am J Sociol 1973 May;78(6):1360-1380. [doi: 10.1086/225469]
19. Berkman LF, Syme SL. Social networks, host resistance, and mortality: a nine-year follow-up study of Alameda County residents. Am J Epidemiol 1979;109(2):186-204 [FREE Full text] [doi: 10.1093/oxfordjournals.aje.a112674] [Medline: 425958]
20. House J. Structures and processes of social support. Annu Rev Sociol 1988;14:293-318 [FREE Full text] [doi: 10.1146/annurev.soc.14.1.293]
21. Cohen S, Mermelstein R, Kamarck T, Hoberman HM. Measuring the functional components of social support. In: Sarason IG, Sarason BR, editors. Social Support: Theory, Research and Applications. Netherlands: Dordrecht: Springer Netherlands; 1985:73-94.
22. Cohen S. Psychosocial models of the role of social support in the etiology of physical disease. Health Psychol 1988;7(3):269-297. [doi: 10.1037/0278-6133.7.3.269] [Medline: 3289916]
23. Cohen S, Wills TA. Stress, social support, and the buffering hypothesis. Psychol Bull 1985 Sep;98(2):310-357. [doi: 10.1037/0033-2909.98.2.310] [Medline: 3901065]
24. Thoits PA. Social support as coping assistance. J Consult Clin Psychol 1986;54(4):416-423 [FREE Full text] [doi: 10.1037//0022-006x.54.4.416] [Medline: 3745593]
25. Bolger N, Zuckerman A, Kessler RC. Invisible support and adjustment to stress. J Pers Soc Psychol 2000 Dec;79(6):953-961. [doi: 10.1037/0022-3514.79.6.953] [Medline: 11138764]
26. Uchino BN, Bowen K, de Grey RK, Mikal JP, Fisher EB. Social support and physical health: models, mechanisms, and opportunities. In: Fisher EB, Cameron LD, Christensen AJ, Ehlert U, Guo Y, Oldenburg B, et al, editors. Principles and Concepts of Behavioral Medicine: A Global Handbook. Germany: Springer; 2018:341-372.
27. Allen JD, Savadatti S, Levy AG. The transition from breast cancer 'patient' to 'survivor'. Psychooncology 2009 Jan;18(1):71-78. [doi: 10.1002/pon.1380] [Medline: 18613299]
28. Black EK, White CA. Fear of recurrence, sense of coherence and posttraumatic stress disorder in haematological cancer survivors. Psychooncology 2005 Jun;14(6):510-515 [FREE Full text] [doi: 10.1002/pon.894] [Medline: 15669018]
29. Muscatell KA, Eisenberger NI, Dutcher JM, Cole SW, Bower JE. Links between inflammation, amygdala reactivity, and social support in breast cancer survivors. Brain Behav Immun 2016 Mar;53:34-38 [FREE Full text] [doi: 10.1016/j.bbi.2015.09.008] [Medline: 26384778]
30. Hughes S, Jaremka LM, Alfano CM, Glaser R, Povoski SP, Lipari AM, et al. Social support predicts inflammation, pain, and depressive symptoms: longitudinal relationships among breast cancer survivors. Psychoneuroendocrinology 2014 Apr;42:38-44 [FREE Full text] [doi: 10.1016/j.psyneuen.2013.12.016] [Medline: 24636499]
31. Kroenke CH, Kwan ML, Neugut AI, Ergas IJ, Wright JD, Caan BJ, et al. Social networks, social support mechanisms, and quality of life after breast cancer diagnosis. Breast Cancer Res Treat 2013 Jun;139(2):515-527 [FREE Full text] [doi: 10.1007/s10549-013-2477-2] [Medline: 23657404]
32. Salonen P, Tarkka MT, Kellokumpu-Lehtinen PL, Koivisto AM, Aalto P, Kaunonen M. Effect of social support on changes in quality of life in early breast cancer patients: a longitudinal study. Scand J Caring Sci 2013;27:396-405 [FREE Full text] [doi: 10.1111/j.1471-6712] [Medline: 22834764]
33. McDonough MH, Sabiston CM, Wrosch C. Predicting changes in posttraumatic growth and subjective well-being among breast cancer survivors: the role of social support and stress. Psychooncology 2014 Jan;23(1):114-120. [doi: 10.1002/pon.3380] [Medline: 23983131]

34. Fong AJ, Scarapicchia TM, McDonough MH, Wrosch C, Sabiston CM. Changes in social support predict emotional well-being in breast cancer survivors. Psychooncology 2017 May;26(5):664-671. [doi: 10.1002/pon.4064] [Medline: 26818101]
35. Wortman CB. Social support and the cancer patient. Conceptual and methodologic issues. Cancer 1984 May 15;53(10 Suppl):2339-2362 [FREE Full text] [doi: 10.1002/cncr.1984.53.s10.2339] [Medline: 6367944]
36. Helgeson VS, Cohen S. Social support and adjustment to cancer: reconciling descriptive, correlational, and intervention research. Health Psychol 1996 Mar;15(2):135-148. [doi: 10.1037//0278-6133.15.2.135] [Medline: 8681922]
37. Perrin A, Kumar M. Pew Research Center. 2019. About Three-in-Ten US Adults Say They Are 'Almost Constantly' Online URL: https://www.pewresearch.org/fact-tank/2019/07/25/americans-going-online-almost-constantly/ [accessed 2019-09-03]
38. Rainie L, Wellman B. Networked: The New Social Operating System. Boston, MA: The MIT Press; 2012.
39. Wellman B, Haase AQ, Witte J, Hampton K. Does the internet increase, decrease, or supplement social capital? Am Behav Sci 2016 Jul 27;45(3):436-455. [doi: 10.1177/00027640121957286]
40. Rains SA, Keating DM. The social dimension of blogging about health: health blogging, social support, and well-being. Commun Monogr 2011 Dec;78(4):511-534 [FREE Full text] [doi: 10.1080/03637751.2011.618142]
41. Walther JB, Boyd S. Attraction to computer-mediated social support. In: Lin CA, Atkin DJ, editors. Communication Technology and Social Change: Theory and Implications. Newburyport, Massachusetts: Hampton Press; 2002:153-188.
42. Lieberman MA, Golant M, Giese-Davis J, Winzlenberg A, Benjamin H, Humphreys K, et al. Electronic support groups for breast carcinoma. Cancer 2003 Feb 3;97(4):920-925 [FREE Full text] [doi: 10.1002/cncr.11145] [Medline: 12569591]
43. Kornblith AB, Herndon JE, Zuckerman E, Viscoli CM, Horwitz RI, Cooper MR, Cancer and Leukemia Group B. Social support as a buffer to the psychological impact of stressful life events in women with breast cancer. Cancer 2001 Jan 15;91(2):443-454. [doi: 10.1002/1097-0142(20010115)91:2<443::aid-cncr1020>3.0.co;2-z] [Medline: 11180093]
44. Moon T, Chih M, Shah DV, Yoo W, Gustafson DH. Breast cancer survivors' contribution to psychosocial adjustment of newly diagnosed breast cancer patients in a computer-mediated social support group. Journalism Mass Commun Q 2017 Jan 19;94(2):486-514 [FREE Full text] [doi: 10.1177/1077699016687724]
45. Namkoong K, Shah DV, Gustafson DH. Offline social relationships and online cancer communication: effects of social and family support on online social network building. Health Commun 2016 Nov 8;32(11):1422-1429. [doi: 10.1080/10410236.2016.1230808] [Medline: 27824257]
46. Rehse B, Pukrop R. Effects of psychosocial interventions on quality of life in adult cancer patients: meta analysis of 37 published controlled outcome studies. Patient Educ Couns 2003 Jun;50(2):179-186. [Medline: 12781933]
47. Attai DJ, Cowher MS, Al-Hamadani M, Schoger JM, Staley AC, Landercasper J. Twitter social media is an effective tool for breast cancer patient education and support: patient-reported outcomes by survey. J Med Internet Res 2015 Jul 30;17(7):e188 [FREE Full text] [doi: 10.2196/jmir.4721] [Medline: 26228234]
48. Hargreaves S, Bath PA, Duffin S, Ellis J. Sharing and empathy in digital spaces: qualitative study of online health forums for breast cancer and motor neuron disease (amyotrophic lateral sclerosis). J Med Internet Res 2018 Jun 14;20(6):e222 [FREE Full text] [doi: 10.2196/jmir.9709] [Medline: 29903695]
49. Sandaunet AG. The challenge of fitting in: non-participation and withdrawal from an online self-help group for breast cancer patients. Sociol Health Illn 2008 Jan;30(1):131-144 [FREE Full text] [doi: 10.1111/j.1467-9566.2007.01041.x] [Medline: 18254837]
50. Cabling ML, Turner JW, Hurtado-de-Mendoza A, Zhang Y, Jiang X, Drago F, et al. Sentiment analysis of an online breast cancer support group: communicating about tamoxifen. Health Commun 2017 Jul 5;33(9):1158-1165 [FREE Full text] [doi: 10.1080/10410236.2017.1339370] [Medline: 28678549]
51. Zhang R. The stress-buffering effect of self-disclosure on Facebook: an examination of stressful life events, social support, and mental health among college students. Comput Hum Behav 2017 Oct;75:527-537. [doi: 10.1016/j.chb.2017.05.043]
52. Vogel EA, Rose JP, Crane C. 'Transformation tuesday': temporal context and post valence influence the provision of social support on social media. J Soc Psychol 2017 Nov 10;158(4):446-459 [FREE Full text] [doi: 10.1080/00224545.2017.1385444] [Medline: 29023225]
53. Klemm P, Bunnell D, Cullen M, Soneji R, Gibbons P, Holecek A. Online cancer support groups: a review of the research literature. Comput Inform Nurs 2003;21(3):136-142. [doi: 10.1097/00024665-200305000-00010] [Medline: 12792194]
54. Edosomwan S, Prakasan SK, Kouame D, Watson J, Seymour T. The history of social media and its impact on business. J Appl Manag Entrepreneursh 2011;16(3):79-91 [FREE Full text]
55. Snelson CL. Qualitative and mixed methods social media research. Int J Qual Methods 2016 Mar 4;15(1):160940691562457. [doi: 10.1177/1609406915624574]

Edited by G Eysenbach; submitted 27.11.18; peer-reviewed by W Tsai, HM Tsai, D Giustini, T Rashid Soron; comments to author 31.03.19; revised version received 27.05.19; accepted 21.07.19; published 24.10.19.

Please cite as:
Mikal JP, Grande SW, Beckstrand MJ
Codifying Online Social Support for Breast Cancer Patients: Retrospective Qualitative Assessment
J Med Internet Res 2019;21(10):e12880
URL: http://www.jmir.org/2019/10/e12880/
doi:10.2196/12880
PMID:31651404

Original Paper

Difficulties Encountered by People With Depression and Anxiety on the Web: Qualitative Study and Web-Based Expert Survey

Renaldo Bernard[1], MSc (Psych), MSc (WebSci); Carla Sabariego[1], PhD; Alarcos Cieza[2], PhD

[1]Chair for Public Health and Health Services Research, Research Unit for Biopsychosocial Health, Department of Medical Informatics, Biometry and Epidemiology, Ludwig-Maximilians-Universität München, Munich, Germany

[2]Blindness and Deafness Prevention, Disability and Rehabilitation, World Health Organization, Geneva, Switzerland

Corresponding Author:
Renaldo Bernard, MSc (Psych), MSc (WebSci)
Chair for Public Health and Health Services Research, Research Unit for Biopsychosocial Health
Department of Medical Informatics, Biometry and Epidemiology
Ludwig-Maximilians-Universität München
Marchioninistraße 17
Munich, 81377
Germany
Phone: 49 49 89 2180 7822
Email: renaldo.bernard@med.lmu.de

Abstract

Background: Depression and anxiety are the most common mental health conditions, and they were identified as leading contributors to global disability in 2016. People with these conditions rely on Web-based resources as a source of accurate health information, convenient and effective treatment, and essential social support. However, a recent systematic review revealed several potentially limiting difficulties that this group experiences online and also suggested that there is a partial understanding of these difficulties as only difficulties associated with neurocognitive, but not sociocognitive, deficits were identified. Therefore, this study fills this knowledge gap and contributes to a more robust and fuller understanding of the difficulties this group experiences online.

Objective: The objective of this study was to identify the difficulties people with depression and anxiety experience when using the Web and the Web activities that are most associated with the experience of difficulties.

Methods: The study employed data triangulation using face-to-face semistructured interviews with 21 participants affected by depression and anxiety and a comparison group (7 participants) without mental disorders (study 1) as well as a persona-based expert online survey with 21 mental health practitioners (MHPs) who treated people with depression and anxiety (study 2). Framework analysis for both studies proceeded through 5 stages: (1) familiarization, (2) identifying a thematic framework, (3) indexing, (4) charting, and (5) mapping and interpretation.

Results: In study 1, 167 difficulties were identified from the experiences of participants in the depression and anxiety group were discussed within the context of 81 Web activities, services, and features. From these, 4 themes and 12 subthemes describing the difficulties people with depression and anxiety experienced online were identified. Difficulties relating to the subtheme lack of control over access and usage were the most common difficulties experienced by participants in the depression and anxiety group (19/21). Sixteen difficulties identified from the experiences of participants in the comparison group were discussed within the context of 11 Web activities, services, and features. Most participants in the comparison group (6/7) contributed to the subtheme describing difficulties with unexpected and irrelevant content. In study 2, researchers identified 3 themes and 10 subthemes that described the perceived difficulties people with depression and anxiety might experience online as reported by MHPs. Practitioners linked these difficulties with 22 common impairments, limitations in activities of daily life, and diagnostic criteria associated with depression and anxiety.

Conclusions: People with depression and anxiety also experience difficulties when using the Web that are related to the sociocognitive deficits associated with their conditions. MHPs have a good awareness of the difficulties that people with depression and anxiety are likely to experience when using the Web. This investigation has contributed to a fuller understanding of these difficulties and provides innovative guidance on how to remove and reduce them for people with depression and anxiety when using the Web.

International Registered Report Identifier (IRRID): RR2-10.1007/978-3-319-21006-3_3

(J Med Internet Res 2019;21(10):e12514) doi:10.2196/12514

KEYWORDS

World Wide Web; depression; anxiety; accessibility; interview; persona; expert study; eHealth; usability; user experience; facilitators; barriers; mental disorders

Introduction

Background

Depression and anxiety are the most common mental health conditions, and they were identified as leading contributors to global disability in 2016 [1]. Among other resources, people with depression and anxiety often rely on the Web as a resource for health information gathering [2,3], a source of convenient and effective treatment across the life span [4,5], and as a means to connect with others and receive social support [6].

However, a recent systematic review revealed that people with mental health conditions experience difficulties when using the Web [7] that might limit how much they benefit from the Web and that are also poorly understood. The review highlighted a narrow range of difficulties solely related to neurocognitive dysfunction, that is, impaired attention, processing and responding to information slowly, and problem solving. Although sociocognitive deficits, that is, impaired affect regulation and difficulty processing emotional cues, are also as important features of mental health conditions as neurocognitive deficits [8,9], no difficulties relating to sociocognitive deficits were identified by the 13 included studies. The review suggests that despite a relatively recent surge of interest within the general field of human-computer interaction into sociocognitive phenomena [10], this trend seemingly does not apply to Web accessibility research focused on people with mental health conditions. For example, the included studies did not employ methods that would unearth possible difficulties related to sociocognitive deficits, and other researchers have arrived at similar conclusions [11].

Therefore, this investigation will provide a broader perspective that could fill the abovementioned knowledge gap and contribute to a more robust and fuller understanding of the difficulties people with depression and anxiety experience when using the Web. This knowledge will primarily assist Web professionals in creating more accommodating experiences for people with depression and anxiety online and allow this group better access to the opportunities available to everyone else using the Web. It is also expected that the Web would benefit from greater inclusivity where people with depression and anxiety could also make valuable contributions to this informational resource. This research was conducted under the BETTER (weB accEssibiliTy for people wiTh mEntal disoRders) project that investigates Web accessibility for people with depression and anxiety and focuses on these conditions because of their high burden relative to that of other mental health conditions [12].

Objectives

The objective of this study was to identify the difficulties people with depression and anxiety experience when using the Web and the Web activities for which the most difficulties are reported. This study specifically aimed to achieve its objective through triangulation [13] using 2 data sources: face-to-face interviews with people with depression and anxiety and a comparison group without mental disorders (study 1) as well as a persona-based expert Web-based survey with mental health practitioners (MHPs) who treated people with depression and anxiety (study 2).

Methods

Study Designs

Semistructured interviews with people with depression and anxiety and a comparison group were conducted in study 1, and an MHP expert online survey was conducted in study 2. Data triangulation [13] was used to add breadth and depth to the analysis and to evaluate the robustness of findings. Although useful for strengthening the study's conclusions and reducing the risk of false interpretations, triangulation was not used to reduce the findings to a single common truth or for validating one view with another view. Of all the stakeholders in the care of people with depression and anxiety, the person themselves and their therapists were the most accessible and well-suited sources of insight available to this investigation. Study 1 and 2 data were collected, aggregated, and later discussed.

Study 1—A Semistructured Interview Study With a Comparison Group

Recruitment

Participants were recruited using purposive sampling, first with the aim of maximum variation [14,15], based on age [16,17], gender [18], and condition severity [19] as these factors have been found to influence Web usage. However, as this recruitment strategy proved difficult to acquire participants overtime, potential participants who met the inclusion criteria were later considered for participation as well. Participants were recruited using posters placed on message boards around the university and short recruitment messages broadcasted via the university's intranet news feed and social media accounts. Participants were recruited in the United Kingdom based on the following criteria: they were aged ≥18 years, skilled Web users, diagnosed with depression and or an anxiety disorder, and had no sensory or physical impairments that required the use of adaptive or assistive technologies to operate a computer system. It was important only to include skilled Web users to reduce the likelihood that difficulties encountered could be associated with being an unskilled user rather than the targeted conditions. Participants in the comparison group were recruited to match the demographic profile of participants with a diagnosis, but with 1 exception, these participants were encouraged by recruitment advertisements and the participant information sheet to only consider participating if they were never diagnosed with a mental disorder. Data saturation (ie, no new data, themes, and coding) [20] determined the final number of participants to

recruit. Data saturation helps to ensure that the study is supported by adequate and quality data [21].

Screening tools were used to ensure that participants met the inclusion and exclusion criteria. Potential participants in both groups scoring more than 25 on the 10-item abbreviated Web use skills index for the general population [22] were invited to participate. Beck Depression Inventory-II [23] and Beck Anxiety Inventory [24] measure symptom severity at 3 levels (ie, mild, moderate, and severe) and were used for this purpose. Those in the comparison group were assessed for depression and anxiety using the Patient Health Questionnaire for Depression and Anxiety (score between 0 and 2), which has demonstrated high sensitivity and specificity in screening for both conditions [25] and has much fewer items than the other 2 instruments.

Data Collection

Ethical approval was granted by the ethics committee of the University of Southampton. Those who passed screening and gave written consent were invited to participate in a face-to-face semistructured interview lasting between 60 and 90 min. Semistructured interviews allowed researchers to gather rich descriptive data about the experiences of participants when using the Web. The method is also useful for exploring this research domain that is in its infancy [26]. Furthermore, it allows for the flexibility to pursue unexpected experiential paths as shared by the participants without losing focus on the key issues of investigation [26]. A topic guide (Multimedia Appendix 1) was used to ask questions about the difficulties participants experienced when using the Web during their daily lives. The interviews were conducted in private rooms around the university between June and November 2016 and were transcribed verbatim from digital audio recordings and evaluated for accuracy before being analyzed. Personally identifiable data were removed from transcriptions, and pseudonyms were used for participants.

Data Analysis

Framework analysis as outlined by Ritchie et al [27] is commonly used to analyze stakeholder accounts from in-depth semistructured interviews. Although the technique primarily subscribes to a thematic approach, it also permits identified themes from semistructured interview narratives to be organized around research questions [28]. Aided by VERBI's MAXQDA 12 qualitative research software package, researchers proceeded through 5 stages: (1) familiarization, (2) identifying a thematic framework, (3) indexing, (4) charting, and (5) mapping and interpretation.

Researchers became familiar with collected data by listening to the recordings and reading and rereading transcripts while progressively making initial notes of any thoughts that surfaced. Themes were then identified and questioned. Data were sifted, and selected quotes were sorted and later rearranged thematically [29]. The discovered themes were compared to ensure they accurately reflected the data. The analysis then went into a deeper interpretative phase focusing on extracts that illuminated participants' accounts in vivid detail.

Study 2—A Mental Health Practitioner Expert Online Survey Study

Recruitment

Purposive sampling, specifically expert sampling [30,31], was used to recruit participants for this survey. Respondents had to be aged ≥18 years; had to be an accredited, a chartered, or a registered member of a professional body in the United Kingdom for MHPs; and must have had experience treating people with depression and anxiety. MHPs were considered suitable experts for this study as they aim to improve their patients' mental health through therapy that benefits from a deep understanding of their patients' lived experiences [32]. Therefore, it was expected that MHPs would have a good understanding of the difficulties this group might encounter online as well. We examined the potential difficulties people with depression and anxiety face on the Web, as explained by MHPs included in an online database directory of MHPs between January and October 2016. Data saturation [20] determined the final number of respondents. The Checklist for Reporting Results of Internet E-Surveys [33] for this survey is presented in Multimedia Appendix 2.

Data Collection

Ethical approval was granted by the ethics committee of the University of Southampton. Respondents gave their consent before participating in the Web-based survey, which was conducted between January and October 2016. They then answered questions relating to 2 of the 4 personas that were randomly given. One persona focused on depression and the other on an anxiety disorder.

The survey asked demographic questions (eg, educational background and expertise) and open-ended questions about the personas that were provided (Multimedia Appendix 3). The 4 personas used were fictional characters (2 with depression and 2 with anxiety) developed for this study by RB based on information about impairments, activity limitations and participation restrictions experienced by people with depression and anxiety, diagnostic criteria associated with these conditions, and also scenarios that featured a wide range of common Web activities.

The International Classification of Functioning, Disability, and Health Core Set for Depression is an internationally accepted and evidence-based selection of functioning domains [34] that covers the spectrum of symptoms and limitations in the functioning of persons with depression. The seminal study on horizontal epidemiology [35] involving systematic literature reviews, content analysis of patient-reported outcomes and outcome instruments, clinical input, and a qualitative study generated a useful group of psychosocial difficulties commonly experienced across brain disorders. The International Statistical Classification of Diseases and Related Health Problems 10th edition is a classification created and maintained by the World Health Organization [36]. The Web activity taxonomy developed by Sellen et al [37], and that later received strong support from Kellar et al [38], was used to animate personas presented to respondents. Feedback on the first version of the survey was

obtained from 2 MHPs who participated in the survey pilot, and some modifications were made as a result.

Data Analysis

The data analysis technique used on the data from the semistructured interviews was also applied to these data.

Results

Summary of Results

A total of 167 difficulties that people with depression and anxiety experienced when using the Web were identified in study 1, and 10 difficulties were identified in study 2. A comparison of these findings will be shared after the findings for each study are detailed below.

Study 1—Semistructured Interview Study With a Comparison Group

This study had a sample of 28 participants (8 males and 20 females) aged between 18 and 69 years (Table 1). A total of 16 females and 5 males were recruited for the sample of people with depression and anxiety. Moreover, 4 females and 3 males were included in the comparison group. An independent samples *t* test was conducted to compare the level of Web skill in the depression and anxiety group with that in the comparison group. There was no significant difference in the level of Web skill in the depression and anxiety (mean 40.10, SD 6.94) group and comparison (mean 41.14, SD 9.26) group (t_{26}=0.32; *P*=.75, 2-tailed). A comparison of the identified difficulties encountered by participants in both groups reveals that they have noticeably different experiences when using the Web.

Table 1. Sample demographics of the semistructured interview study with a comparison group.

Characteristics	Depression and anxiety group (n=21), n	Comparison group (n=7), n
Sex		
Male	5	3
Female	16	4
Age (years)		
18-29	11	5
30-49	7	1
50-69	3	1
Condition		
Anxiety	2	—[a]
Depression	9	—
Anxiety and depression	10	—
None	—	7
Condition severity: depression		
Mild	2	—
Moderate	6	—
Severe	11	—
Condition severity: anxiety		
Mild	0	—
Moderate	5	—
Severe	4	—
Frequency of Web usage		
Several times a day	20	6
Once a day	1	1
Several times a week	0	0
Once a week	0	0
Once a month or less	0	0

[a]Not applicable.

The 167 difficulties identified from the experiences of participants in the depression and anxiety group were discussed within the context of 81 Web activities, services, and features. The majority of participants in the depression and anxiety group reported difficult experiences that were captured in each theme (Table 2).

The 16 difficulties identified from the experiences of participants in the comparison group were discussed within the context of 11 Web activities, services, and features. These difficulties were often encountered with momentary negative affect and, occasionally, resulted in dislike for the particular Web activity, service, or feature. Difficult experiences shared by most participants in the comparison group were represented by theme 1 (Table 2). Each of the remaining 3 themes included difficult experiences discussed by a small number of participants in this group (Table 2).

Table 2. Number of participants in the people with depression and anxiety group compared with number of participants in the comparison group by theme.

Theme	Participants in the people with depression and anxiety group (n=21), n	Participants in the comparison group (n=7), n
Inappropriate and sensitive content	20	6
Lack of safety, privacy, and security controls	20	4
Lack of adequate support	19	1
Difficult user interfaces	17	3

Theme 1: Inappropriate and Sensitive Content

A total of 4 subthemes were identified within this theme: (1) unexpected, irrelevant, and inappropriate content is upsetting; (2) reminders of upsetting experiences and negative affect triggers; (3) social comparison cues on social media that result in increased negative affect; and (4) abusive content limits Web usage by those who avoid it. The majority of participants, 20 out of 21 in the depression and anxiety group and 6 out of 7 in the comparison group, identified inappropriate content as a source of difficulty under this theme. Other subthemes were only reported by participants in the depression and anxiety group who discussed difficulties with sensitive and abusive content (Table 3).

Unexpected, Irrelevant, and Inappropriate Content Is Upsetting

Exposure to inappropriate content was followed by feelings of upset, frustration, and helplessness. This negatively impacted the ability of some participants in the depression and anxiety group to complete tasks for up to a day:

> *See a photo and it's affected my mood for the rest of the day. I'll be there sat when I go to work, flip through social media, all of a sudden something hits, feel low, go to work and it doesn't pick up and then I can't perform at work and then I get sent home which makes me feel even worse.* [Shane, depression and anxiety group]

Participants in the comparison group generally felt upset by inappropriate content but saw such content as being a regular part of using the Web and appeared better able to quickly overcome these feelings than those in the depression and anxiety group:

> *Was it upsetting for a long time or was it just that moment?* [Moderator]
>
> *Just that initial moment.* [Marita, comparison group]

Reminders of Upsetting Experiences and Negative Affect Triggers

Exposure to sensitive content resulted in involuntarily recollecting memories of personally meaningful issues that were upsetting for a temporary or prolonged period.

Such sensitive content on the Web is varied and diverse, as shown in Table 3. Some types of sensitive content are also composed to provoke a strong negative emotional response deliberately, for example, graphic content used in news stories and promoted posts on social media related to appeals by causes supporting people and animals in need:

> *It's almost like some adverts I can't watch because I just think, "I mean, I know they are poor, starving children in Africa"...I'm paying you know, and I'm doing [my] bit. But I literally get to the point, I sit and think, "Oh my God, if that was me, if that was my child," I mean, I would just give away, I could never, I'd just be giving away my clothes.* [Clara, depression and anxiety group]

No participant in the comparison group reported a similar difficulty.

Table 3. Difficulties with sensitive and inappropriate content experienced by participants in both sample groups by subtheme.

Subtheme—number of participants from DA[a] (n=21) and comparison (n=7) groups	Difficulty reported by the DA group participants	Difficulty reported by the comparison group participants
Unexpected, irrelevant, and inappropriate content is upsetting (DA group: 17/21 and comparison group: 6/7)	• Exposure to upsetting offensive content from social connections (eg, violence, trifle, overshare, exaggeration, and constant help seeking), news websites (eg, violence, headline marquee, articles, and political bias), and advertising (eg, prominently positioned, excessive amounts on page, disguised, misleading, obstructive, persistent, distracting, and intrusive) • Notifications highlighting insignificant information on social media platforms • Unexpected and inappropriate search results	• Exposure to offensive content and personally critical comments from social connections, online dating counterparts (eg, sexual content and inappropriate contact), news websites (eg, untrustworthy articles, political bias, and violent acts), and advertising • Notifications from social media platforms
Reminders of upsetting experiences and negative affect triggers (DA group: 14/21 and comparison group: 0/7)	• Inappropriate help-seeking behavior on social media by those with similar negative experiences • Social media features—highlighting content such as status updates, images, and posts from social connections from current date in the past—that trigger memories of upsetting experiences • Personally relevant content (eg, status updates, images, posts from friends, adverts, and news articles) that triggers negative affect	—[b]
Social comparison cues on social media that result in increased negative affect (DA group: 7/21 and comparison group: 0/7)	• Social media content (eg, images, information on healthy lifestyle practices, and past positive life experiences) highlighting perceived personal faults • Social media content (eg, images) highlighting opportunities that are no longer available to one's self but to others who are similar • Instructive content, especially user-generated, that is related to sensitive topics (eg, child-rearing) and that is contrary to personal practices	—
Abusive content limits Web usage by those who avoid it (DA group: 5/21 and comparison group: 0/7)	• Avoidance of unfamiliar and news-related websites because of the fear of unintentionally accessing personally upsetting and inappropriate content • Avoidance of social media participation because of the fear of receiving abuse	—

[a]DA: depression and anxiety.

[b]Not applicable.

Social Comparison Cues on Social Media That Result in Increased Negative Affect

Some content was effective at directing attention toward drawing upward comparisons between several participants in the depression and anxiety group and others, and participants in the depression and anxiety group and themselves in past. These comparisons were often negative and considerably upsetting:

> *I had one that came up this week that was a photograph of me, many years ago. Friend's wedding. I was a bridesmaid. I just looked at this photograph and went, I mean, I looked good…I immediately felt that I'd let myself down. I thought, "Well look, clearly you can manage this. What's happened?"* [Clara, depression and anxiety group]

No participant in the comparison group reported a similar difficulty.

Abusive Content Limits Web Usage by Those Who Avoid It

Some participants also refrained from commenting, posting updates, and engaging in various Web-based activities in fear of suffering abuse from other Web users as a result:

> *I left a comment, and then I just had a stream of abuse from people, because I voted to leave... The Web, in general, is quite a hostile place, and I don't want to be in that sort of environment. It doesn't make me feel particularly safe or comfortable, being online. As I said, I stick very much to what I know, because I feel quite unsafe outside of that.* [Jason, depression and anxiety group]

No participant in the comparison group reported a similar difficulty.

Theme 2: Lack of Safety, Privacy, and Security Controls

A total of 4 subthemes were identified within this theme: (1) lack of control over access and usage, (2) lack of safety controls, (3) threats to privacy, and (4) ambivalent contact. Ensuring safety for oneself and significant others when using the Web was described as a difficult task only by participants in the depression and anxiety group, as shown in Table 4. Similarly, except for 1 participant from the comparison group, difficulties pertaining to contact were reported by several participants in the depression and anxiety group (10/21).

Lack of Control Over Access and Usage

Table 4 shows several recreational activities on the Web that several participants (19/21) from the depression and anxiety group said displaced important tasks. Participants were unable to stop engaging in these activities even when they wanted to stop. Some participants said that these activities were an outlet for coping with unpleasant feelings and procrastination:

> *I discovered YouTube over the winter exams…When I'm overloaded in other areas, it's like a release…I know I'm doing it. That does not mean I can stop.* [Sara, depression and anxiety group]

A high-level of ease of use was also attributed to making unintended purchasing, banking, and time management decisions, without giving due consideration. These features were considered as being too easy not to use:

> *There's the one-click, it's so easy just to go through and buy and buy and buy, and buy loads of stuff that you can't really afford…Like I said, I have a tendency, sometimes, to make impulse purchases, and I'll look and think I've got more money than I have, and before I know it I'm at the bottom of my overdraft again.* [Jason, depression and anxiety group]

Lack of Safety Controls

Some participants discussed how they grappled with complex issues relating to the differences in privacy approaches between countries and companies and the repercussions for what they self-disclosed to websites based on these factors. Other participants were concerned about keeping their children safe but admitted that they were unable to remain motivated to keep abreast of the constant changes in how safety was managed and circumvented on various websites. Several participants expressed their interest in realizing the wider benefits of the Web. However, they were forced to strictly limit their use of many websites, such as social media websites, and others narrowed their use of the Web to a limited number of websites in fear of abuse and receiving unsolicited contact.

The fear of being a victim of crime and getting involved in a conflict on the Web is equally as concerning as avoiding abuse and unsolicited contact. The result of dealing with this fear is often also limiting Web use. A lack of forewarning about the known types of service misuse, information on how to avoid safety pitfalls associated with the usage of Web-based services and a lack of support options in the event the user is negatively affected also presented difficulties for depression and anxiety group.

As demonstrated with theme 1, it is important that participants have a choice in what content they are exposed to, especially on social media platforms, as the emotional consequences can be profound. Participants were exposed to sensitive content regularly and were unable to avoid it effectively. The highly varied nature of sensitive content on the Web and the lack of control over exposure to it were the main reasons given for why this occurred:

> *People post videos of the dogs being boiled alive to raise awareness...It's a really upsetting video, you don't have the possibility to not want to play it, you got auto play on and you scroll through it, it would just start playing.* [Jade, depression and anxiety group]

No participant in the comparison group reported a similar difficulty.

Threats to Privacy

Participants in both groups were generally concerned about the privacy of their personal data. Participants in the depression and anxiety group identified many instances of where they particularly felt vulnerable, as shown in Table 5.

However, these participants sometimes also failed to take necessary precautions because of their felt sense of personal insignificance:

> *In terms of difficulties, it's really kind of finding minor details for terms and conditions for various services and various things that you use online, whether it's the rights that a social media platform has for your data or the rights of a purchasing website to then use your details in marketing. It's very buried, I find.* [Betty, depression and anxiety group]

Ambivalent Contact

Several participants (10/21) in the depression and anxiety group experienced much difficulty with direct contact from social connections and stopping consistent contact from unknown senders:

> *My partner almost caused me to lose my life…I don't follow him, I'm not friends with him. And then suddenly, about two weeks ago. On the bottom of photo, he wrote something…That's really unsettling.* [Hera, depression and anxiety group]

> *If somebody messaged me personally I would always respond...I think it gets worse when I'm low...I find social interactions quite draining, when I'm already tired, because you kind of in a way have to put up a bit of a façade, which is obviously very hard to maintain.* [Paisley, depression and anxiety group]

Moreover, 1 out of 7 participants in the comparison group identified avoiding spam via social media platforms as being a difficult task.

Table 4. Difficulties because of a lack of safety, privacy, and security controls experienced by participants in both sample groups by subtheme.

Subtheme—number of participants from DA[a] (n=21) and comparison (n=7) groups	Difficulty reported by the DA group participants	Difficulty reported by the comparison group participants
Lack of control over access and usage (DA group: 19/21 and comparison group: 1/7)	• Addictively accessing similarly upsetting content (eg, news articles on similar topics) that is readily available • Repeatedly clicking on posts and performing other actions on social media, news, and shopping websites • Addictively performing online tasks that displace other tasks—gaming, gambling, and watching videos • Keeping track of time on social media is difficult • Coping with anxiety by fixating on finding answers to a salient issue online and avoiding activities (eg, accessing bank account in anticipation of a low balance) • Easy achievable compulsion to set up bank overdrafts, make online purchases, and donate to charities • Avoid online civic engagement because of a sense of insignificance • Personal online shopping results in feelings of guilt • Reluctantly using online dating when feeling low to increase feelings of self-worth	• Addictively accessing social networking services and news websites
Lack of safety controls (DA group: 11/21 and comparison group: 0/7)	• Understanding how to protect one's family from online dangers and being confident enough to do so • Unable to anticipate if a news article will be upsetting • Detecting scams and phishing attempts on banking platforms • Limiting Web usage by only using familiar websites to avoid unknowingly committing criminal acts • Lack of control over exposure to content • Trusted websites that occasionally feature links to unsafe websites	—[b]
Threats to privacy (DA group: 10/21 and comparison group: 1/7)	• Fear that data from personal data breach would be sold to third parties, or fear of being hacked • A sense of insignificance discourages the implementation of privacy measures on social media platforms • Frustrating when personal data sharing, including seemingly unnecessary personal data, is required to participate in online activities • Targeted advertising using posts, especially posts shared during a depressive episode • Distressing having comments publicly visible • Finding and understanding terms and conditions policies and keeping abreast of changes	• Ensuring privacy and safety online—identifying scams and scammers
Ambivalent contact (DA group: 10/21 and comparison group: 1/7)	• Fear of direct contact or contact beyond a "*like*" or similar form of engagement, from social connections, especially during a depressive episode • Uncertainty about how to stop contact—being removed from electronic mailing lists • Mandatory contact to obtain resources—subscription to electronic mailing lists • Making contact—connecting with people through video clips and reading news instead of direct contact, avoid responding to messages as it is mentally effortful, and pressured to respond to messages immediately	• Avoiding spam via online social networking sites

[a]DA: depression and anxiety.

[b]Not applicable.

Table 5. Difficulties because of having a lack of adequate support experienced by participants in both sample groups by subtheme.

Subtheme—number of participants from DA[a] (n=21) and comparison (n=7) groups	Difficulty reported by the DA group participants	Difficulty reported by the comparison group participants
Lack of support for error recovery and overcoming emotional difficulties (DA group: 17/21 and comparison group: 3/7)	• Managing subscriptions is frustrating—lack of forewarning, automatic renewals, and difficulty requesting and obtaining refunds • Remembering many passwords and special codes • Lack of clear warnings about the risks associated with online dating on particular platforms and of support when things go wrong • Posting content is mentally effortful and time-consuming—choosing emoji, expressing feelings without causing alarm, and fear of using incorrect grammar • Immediately quit or desperately and hastily try many strategies to complete challenging tasks • Having too many options and information results in indecisiveness and distraction • Impersonal social messaging feature makes taking the first step to seek help from connections difficult • Feeling ignored when social connections do not react to personal posts • Sharing content is a struggle—highlighting personal positives, fear of attracting abuse and attention from others, making offense or causing conflict, and fear of sharing inaccurate and uninteresting content • Difficulty learning from onscreen material because of an inability to engage actively • Difficulty getting support online, given the inflexibility of online banking and education support systems that are often not user-friendly • Unable to accurately gauge reactions when interacting with others via online dating platforms • Fear of opening messages and being pressured to respond to a message immediately as a read receipt has been sent	• Fast time-outs
Information gathering on the Web (DA group: 12/21 and comparison group: 3/7)	• Choosing the right search keywords for difficult-to-find resources • Time-consuming to assess the veracity of information on the Web • Hard to keep focus and understand information online during a depressive episode • Getting distracted when navigating across many websites to find needed information • Search results listing multiple sources with identical information and few sources with original content • Assessing the availability of resources across multiple academic databases • Not knowing when to stop searching for and evaluating online resources • Quit searching the Web in frustration after not finding needed results • Processing too much information from many sources makes it difficult to ascertain if one's search is complete • Easily distracted when browsing the Web during a depressive episode • Sites that break up content across pages to increase advert views	• Selecting the effective Web search terms • Missing information—important information that should be online but is not • Reading multipage articles

[a]DA: depression and anxiety.

Theme 3: Lack of Adequate Support

A total of 2 subthemes were identified within this theme: (1) lack of support for error recovery and overcoming emotional difficulties and (2) information gathering on the Web. The majority of participants (17/21) in the depression and anxiety group and 3 out of 7 participants in the comparison group highlighted that a lack of meaningful support made using the Web challenging (Table 5). Difficulties with information gathering online were also identified by 12 out of 21 participants in the depression and anxiety group and 1 out of 7 participants in the comparison group.

Lack of Support for Error Recovery and Overcoming Emotional Difficulties

Several participants in the depression and anxiety group believed that they were not given the necessary support by websites to overcome difficulties, especially when they were feeling unwell:

> *I make lots of mistakes don't get me wrong. I got delivered five kilos of bananas the other day...They could have a "do you need" button before you submit…"Are you sure you need five kilos or five bananas?"* [Christine, depression and anxiety group]

They also pointed out that websites often compounded this situation by not automatically correcting obvious user errors and, instead, sometimes made completing tasks more effortful as a result. Existing support options were not helpful as they were often not user-friendly and did not address difficulties common to participants. These participants shared that remaining motivated to solve the reported difficulties independently of others was challenging:

> *It depends on how tired I am. If I can't get what I want immediately I give up. Then I'll shelve it, and I'll come back. If I need it really urgently, then I just try lots of different things.* [Hera, depression and anxiety group]

Several participants (3/7) in the comparison group shared difficult experiences with form timeouts that terminated too quickly.

Information Gathering on the Web

Gathering information using the Web, especially via search engines and reading multipage articles, proved challenging for participants in both groups (Table 5):

> *Things that I find difficult are getting that...putting the right stuff into your search...So that you get the information you want, and when you know something's out there, but you can't easily get to it.* [Hera, depression and anxiety group]

> *Then there's one picture on one page then you have to scroll to another page to the next part of the article...I find that really frustrating.* [Christine, depression and anxiety group]

Participants (12/21) in the depression and anxiety group experienced additional difficulties in remaining focused when searching the Web using databases and browsing across multiple websites (Table 5):

> *Where you have to click to go to the next. You know they're just doing that, I feel, to measure their clicks so they know how far you're getting in the story, is what I feel. Especially if I'm just doing it on my phone. I have a cheaper phone. It's not so fast. I think, "Oh okay, forget it." This is annoying when you could just put the content right there one page.* [Kurt, depression and anxiety group]

Theme 4: Difficult User Interfaces

A total of 2 subthemes were identified within this theme: (1) using complicated and effortful user interfaces on the Web and (2) malfunctioning websites. Participants in both groups recounted frustrations using complicated and malfunctioning websites (Table 6).

Using Complicated and Effortful User Interfaces on the Web

Unintuitive websites presented difficulties for participants in both groups (Table 6). Participants (14/21) in the depression and anxiety group recalled experiences involving taking regular breaks and frustratingly struggling until they were able to complete tasks such as reading and shopping online:

> *Why do we have to have pop-ups? It kind of perplexes me why it's so invasive. You just kind of feel like...You almost want to flip channels but you can't.* [Betty, depression and anxiety group]

> *You can sit down with a fixed idea of what you would like and then by looking on the web you've got so many different products...You then pull back from the decision because there's too much to decide from.* [Jason, depression and anxiety group]

Participants in the comparison group (3/7) discussed difficulties completing long website forms and constantly changing user interfaces for frequently used services.

Malfunctioning Websites

Participants in both groups identified difficulties with unresponsive websites, feedback, and page loading errors (Table 6). Experiences with malfunctioning websites sometimes led to catastrophic thinking and a reduced willingness to troubleshoot by those in the depression and anxiety group:

> *I tend to try and avoid going onto my app or looking at my bank statement as much as possible because it makes me really worried. I've actually seen it takes twice as long to log you in so it's almost like the wait and the panic that what little money you've got is taking longer.* [Trish, depression and anxiety group]

Table 6. Difficulties with challenging user interfaces experienced by participants in both sample groups by subtheme.

Subtheme—number of participants from DA[a] (n=21) and comparison (n=7) groups	Difficulty reported by the DA group participants	Difficulty reported by the comparison group participants
Using complicated and effortful user interfaces on the Web (DA group: 14/21 and comparison group: 3/7)	• Unintuitive user interfaces for education on the Web—time-consuming to find course materials, complicated academic databases, and difficult reading via Web-based reading services • Unintuitive user interfaces for banking on the Web—intimidated by terminology and abundance of numbers, unclear system feedback, and setting up new bank recipient is complicated • Relearning user interfaces after changes is difficult, especially when lacking the motivation to explore	• Completing long website forms • Constant user interface changes
Malfunctioning websites (DA group: 11/21 and comparison group: 2/7)	• Sites not optimized for mobile browsing and poor connectivity • Malfunctioning critical website feature delaying completion of an important task • Frustrating to be given options that are not available • Bad video streaming experiences because of poor connectivity • Catastrophic thinking because of malfunction or irregularities in the operation of the website • Remaining motivated to independently resolve complicated problems caused by websites	• Unresponsive websites • Page loading errors

[a]DA: depression and anxiety.

[b]Not applicable.

Summary: The Most Common Difficulties People With Depression and Anxiety Encounter When Using the Web

A total of 8 subthemes describe difficulties experienced by more than half of the participants in the depression and anxiety group (Table 7).

Web Activities, Services, and Features for Which the Highest Number of Difficulties Were Reported

Web activities, services, and features being used by participants when they experienced difficulties were also reported (Table 8). Table 8 shows 19 of these for which a higher number of difficulties were reported than the average number of difficulties reported for a Web activity, service, or feature.

Table 7. Most common difficulties by subthemes with number of participants in both groups affected.

Difficulty subtheme	Participants in the people with depression and anxiety group affected (n=21), n	Participants in the comparison group affected (n=7), n
Lack of control over access and usage	19	1
Unexpected and irrelevant content is upsetting	17	6
Lack of support for error recovery and overcoming emotional difficulties	17	1
Features and content that are reminders of upsetting experiences and negative affect triggers	14	0
Understanding complicated user interfaces on the Web	14	2
Information gathering on the Web	12	2
Malfunctioning websites	11	2
Lack of security controls	11	0
Privacy risks	10	1
Ambivalent contact	10	1
Social comparison cues on social media that result in increased negative affect	7	0
Abusive content limits Web usage by those who avoid it	5	0

Table 8. Web activities, services, and features for which the highest number of difficulties were reported.

Web activities, services, and features	Difficulties reported from the people with depression and anxiety group (n=167), n	Difficulties reported from the comparison group (n=16), n
Facebook	56	3
General Web usage	22	2
News sites	19	0
Adverts	17	0
Online learning	17	0
Online banking	16	2
Online shopping	16	1
Conducting research online	13	0
Content sharing by connections on social networking services	11	1
Posting content	10	0
Business-related Web usage	9	0
Online search	11	0
Online dating	8	0
YouTube	8	0
Twitter	7	1
Online civic engagement	6	0
Commenting feature	6	0
eBay	5	0
Instagram	5	0

Study 2—A Mental Health Practitioner Expert Web-Based Survey Study

Data were collected from 21 respondents (4 males and 17 females) aged between 30 and 72 years using purposive sampling (Table 9).

MHPs reported 10 perceived difficulties relating to Web usage by the people with depression and anxiety. Of these, 2 difficulties were only relevant to personas diagnosed with depression. The remaining 8 difficulties were shared by personas diagnosed with either depression or an anxiety disorder. The 10 perceived difficulties were linked to 22 common impairments, limitations in activities of daily life, and diagnostic criteria associated with depression and anxiety. These difficulties are organized under 3 themes: (1) navigating and generally operating websites, (2) content on the Web, and (3) lack of trust in the Web.

Theme 1: Navigating and Generally Operating Websites

MHPs (n=19) identified 5 perceived difficulties within this theme that focus on the general usage of websites by people with depression and anxiety (Table 10). All 4 personas are captured in this theme.

Navigating the Web

Using the Web was generally seen by experts (n=10) as involving many effortful activities that could pose challenges for how people with depression and anxiety perceived, understood, and used Web-based resources. For example, Web browsing was often pinpointed as a potentially difficult task. Personas with either condition were thought to be lacking the necessary motivation and energy to effectively use Web-based resources and the ability to solve emergent problems. Experts believed that these difficulties could be compounded by impaired emotion regulation, poor concentration, and the physical manifestations of their conditions, for instance, an upset or worried user experiencing difficulty navigating a website along with finding it hard to concentrate on the task at hand.

Malfunctioning and Unintuitive Sites

Malfunctioning websites and websites with an unintuitive design were also thought by experts (n=7) to be especially difficult for people with depression and anxiety to use. These websites were described as using too small font sizes, unnecessarily bright colors, and many shapes within its design. The experts expressed concern that members of this group were often fatigued and might also struggle with remaining resilient when encountering these experiences. Feelings of hopelessness, worthlessness, and worry were mentioned as possible outcomes.

Table 9. Sample demographics of mental health practitioner expert Web-based survey study.

Characteristics	Mental health practitioner experts (n=21), n
Sex	
Male	4
Female	17
Age (years)	
30-49	2
50-69	16
≥70	3
Years of experience	
5-10	3
11-15	5
>15	13
Profession	
Counselor	10
Clinical psychologist	1
Psychiatrist	1
Psychotherapist	8
Occupational psychologist	1
Country	
United Kingdom	20
Ireland	1

Table 10. Perceived difficulties navigating and generally operating websites.

Perceived difficulties	Associated impairments, limitations in activities of daily life, and diagnostic criteria
Navigating the Web[a,b]	Lack of motivation, lack of energy, impaired emotion regulation, poor concentration, physical symptoms (eg, tingling or numb fingers, dizziness, and shortness of breath), and difficulty solving problems
Malfunctioning and unintuitive sites[a,b]	Poor concentration, lack of motivation, low resilience, worry, low mood, low self-confidence, low self-esteem, fatigue, feelings of hopelessness and feelings of worthlessness
Effortful tasks[a,b]	Lack of motivation, worry, impaired emotion regulation, poor concentration and feelings of hopelessness
No clear guidance on how to complete tasks[a]	Poor concentration and feelings of hopelessness
Excessively detailed websites with information/design elements[a,b]	Feelings of being overwhelmed and lack of energy

[a]Difficulty reported for persona with depression.

[b]Difficulty reported for persona with an anxiety disorder.

Effortful Tasks

Several common online activities were highlighted by experts (n=7) as possible areas of difficulty because of the sustained mental effort involved. These activities included seeking help online for technical and personal issues, reading, completing forms, and conducting online research. Experts noted that feelings of hopelessness and worry coupled with a lack of motivation and poor concentration might make these activities challenging for those affected by depression and anxiety.

No Clear Guidance on How to Complete Tasks

The potential for a Web activity to pose difficulty was thought to be increased when no clear guidance on how to complete the necessary tasks was provided. Experts (n=4) believe that this fosters a feeling of hopelessness within users with depression and anxiety.

Excessively Detailed Websites With Information/Design Elements

Some experts (n=3) also characterized websites featuring excessive amounts of information and design elements as being

barriers to effective use. It was feared that information overload would be the likely result and that it would overwhelm people with depression and anxiety who tend to be already low on energy.

Theme 2: Content on the Web

MHPs (n=10) identified 3 perceived difficulties within this theme relating to the perception and comprehension of website content by people with depression and anxiety (Table 11). All 4 personas are captured in this theme.

Table 11. Perceived difficulties with content on the Web.

Perceived difficulties	Associated impairments, limitations in activities of daily life, and diagnostic criteria
Retaining information[a,b]	Poor concentration, worry, fatigue, low self-confidence, and physical symptoms (eg, tingling/numb fingers, dizziness, and shortness of breath)
Content that does not resonate[a,b]	Negativity bias, lack of motivation, impaired emotion regulation, low self-esteem, feelings of being overwhelmed, and feelings of isolation
Content that triggers repetitive thinking[a,b]	Impaired emotion regulation

[a]Difficulty reported for persona with depression.

[b]Difficulty reported for persona with an anxiety disorder.

Retaining Information

The experts (n=4) shared that retaining information on websites would be difficult for users with these conditions. Symptoms including poor concentration, worry, fatigue, and physical symptoms such dizziness were cited as factors that contribute to this outcome.

Content That Does Not Resonate

Experts (n=9) believed that online content lacking personal meaning or importance to users and content that users perceived as overly positive or negative would present several challenges for this group. Experts stated that this kind of unbalanced content could be demotivating, overwhelming, and isolating for this group. Overly negative content was believed to have the added potential to affirm negative fears and concerns. This difficulty was noted for all personas as well:

> *He may have difficulty feeling the wording on a website applies to him if he does not feel directly spoken to in an understanding way by what is written on a website.* [MHP 1]

Content That Triggers Repetitive Thinking

Content that is reminiscent of negative personal experiences was highlighted as a potential challenge by an expert (n=1) for 2 of the personas. Words or phrases that might be associated with these experiences were deemed to have the potential to easily take users on a negative mental tangent where they would repeatedly focus and refocus on negative personal experiences. This kind of repetitive negative thinking [39,40] is believed to present difficulties for concentration and the comprehension of online content as well.

Theme 3: Lack of Trust in the Web

MHPs (n=4) identified 2 perceived difficulties within this theme relating to information sharing by users and their safety online (Table 12). A total of 3 personas—2 diagnosed with depression and 1 an anxiety disorder—are captured in this theme.

Table 12. Perceived difficulties relating to a lack of trust in the Web.

Perceived difficulties	Associated impairments, limitations in activities of daily life, and diagnostic criteria
Self-disclosure online[a,b]	Worry, perfectionism, and low self-confidence
Feeling safe online[a]	Worry, feelings of vulnerability, and withdrawal

[a]Difficulty reported for persona with depression.

[b]Difficulty reported for persona with an anxiety disorder.

Self-Disclosure Online

Sharing personal information online was highlighted as a potential challenge for people in this group. It was mentioned that users might experience great worry about how their information might be used beyond what was intended. Sharing personal information in what seems to be a public setting may also be difficult for users who are experiencing low self-confidence issues. It was also mentioned that some users might worry about making mistakes and become overly concerned about completing information collection forms correctly.

Feeling Safe Online

Ensuring one's safety was also identified as a potential difficulty for people with depression. Fear of privacy violations, exploitation, deception, and crime were some of the issues highlighted. Experts say these fears foster intense feelings of worry and vulnerability that can be mentally debilitating and are believed to result in users withdrawing from some aspects of the Web.

Discussion

Principal Findings and Comparison With Prior Work

This investigation identified and described many of the difficulties people with depression and anxiety experience on the Web through triangulation using 2 data sources. First, study 1 benefited from face-to-face interviews with comparison group participants that were used to highlight unique difficulties identified in face-to-face interviews with participants from the depression and anxiety group. Second, study 2, a persona-based Web-based expert survey, involved MHPs who identified many difficulties that were also reported by the people with depression and anxiety in study 1 and further described these difficulties from a clinical perspective. Findings from this investigation have contributed to a more robust and fuller understanding of the difficulties people with depression and anxiety experience online and provide actionable insight to researchers, engineers, policy makers, and clinicians in their practice.

Participants in the depression and anxiety group reported a higher number of the identified difficulties compared with those in the comparison group. Furthermore, difficulties reported by participants in the depression and anxiety group were discussed within the context of a larger number of Web activities, services, and features when contrasted with the difficulties reported by the participants in the comparison group. In line with research on resilience [41,42], difficulties reported by participants in the depression and anxiety group were more detailed and considered to be more detrimental by those in this group when contrasted with those reported by participants in the comparison group.

Perceived difficulties identified by MHPs in study 2 showed considerable overlap with those identified in study 1 by participants in depression and anxiety group. Nonetheless, the investigation benefited a great deal from using triangulation. People with depression and anxiety clearly detailed the difficulties they experienced, at times also mentioning the harmful consequences, and MHPs suggested linkages between difficulties reported by the people with depression and anxiety and the common impairments, limitations in activities of daily life, and diagnostic criteria associated with their conditions.

Inappropriate and personally sensitive content was instrumental in triggering persistent and cyclic negative thinking that is characteristic of rumination in depression and anxiety [43] and was linked to impaired emotion regulation by MHPs in study 2. This content was also found to encourage negative social comparisons that exacerbated anxiety symptoms and negatively affected daily functioning [44]. Researchers [45] have found that the more Web users with the inclination toward social comparisons engage in this behavior online, the more they experience negative feelings. Findings from some studies [46] also show that exposure to content that supports clear social comparisons is more detrimental to women with a high tendency to compare themselves with others, relative to women without this tendency.

MHPs in study 2 tied the concern of people with depression and anxiety for safety and privacy controls on the Web to a general lack of trust in the Web. Similar to this investigation's findings, research [47] has associated psychological distress with frequent unwanted contact or communications overload because of the usage of multiple online communication channels. Receiving unwanted contact and misunderstandings were identified as the most common negative Facebook experiences reported by young adults in another study [48] and were associated with depressive symptoms as well. The fear of receiving negative feedback online and coping with the pressure to maintain social network updates has also been reported in other studies [49,50]. Moreover, research [51] has also revealed that social anxiety is positively related to a concern for privacy on the Web and that this concern is negatively related to self-disclosure online as well. Safety issues on the Web for people with depression and anxiety also involve distracting features of social media that facilitate procrastination and are likely being used to avoid stressful but necessary tasks [52]. Andreassen and Pallesen [53] found that investing too much time and effort into using social media can negatively impact other social activities, relationships, and well-being. Moreover, despite not being identified by MHP, people with depression and anxiety in other studies [54,55] also shared experiences that involved compulsive buying behavior.

People with depression and anxiety identified many features—regarding making subscriptions easier to manage, issuing notices about known risks on platforms, ability to easily reduce options to choose from, and the ability to have more flexible service support options—that could be implemented to help them overcome many of the difficulties they face when using the Web. Although MHPs did not identify the majority of the difficulties in this theme, these practitioners highlighted difficulties people with depression and anxiety might experience with understanding excessively detailed websites [56] and navigating the Web, which were also mentioned by the people with depression and anxiety.

Both people with depression and anxiety and MHPs identified malfunctioning, unintuitive, and effortful to use sites as descriptions of websites that may pose difficulties to people with depression and anxiety. MHPs suggest the interaction between several characteristics of depression and anxiety (eg, fatigue, poor concentration, and lack of motivation) [9,36] and the aforementioned features of the Web as the reason for the resulting difficulties as described in this paper.

The most common difficulties experienced by participants with depression and anxiety in study 1 were encountered when using the most common Web activities, services, and features mentioned when talking about difficulties. For example, unexpected content is especially common on social networking platforms, news sites, and in online advertisements. Given that the sample had a majority of young participants, the most common Web activities, services, and features where they experienced difficulties were somewhat expected.

Limitations

Though the study benefited from having a mostly young and predominantly female sample in several ways, this may have limited the range of Web activities the sample engaged in and, therefore, also the range of difficulties identified. Similarly, the difference in sample size between the participant groups in study

1 may have also limited the range of difficulties identified for the smaller comparison group. However, data saturation determined the final number of participants for each group, and therefore, no new information that would have enhanced or changed the findings of a study was expected. Nevertheless, this is one of the first studies of its kind, and this population can serve as a good example for future studies with more diverse and larger samples. Moreover, as the small sample sizes in study 1 limit the ability of the independent samples *t* test to detect a statistical difference in Web skills between the depression and anxiety and comparison groups, the results of this analysis should be cautiously interpreted.

Although unlikely, participants in the comparison group could have misreported a past mental disorder diagnosis. However, these participants were also screened using the Patient Health Questionnaire for Depression and Anxiety, and all of them had a score between 0 and 2 at the time of study.

Despite piloting, some MHPs in study 2 deemed the survey as being too long. This sometimes resulted in receiving a few repetitive responses and complaints about the amount of effort necessary to complete the survey. Conducting screening that considers the necessary levels of attentiveness and effort that are needed for such surveys is of utmost importance for future studies involving this group.

Given these limitations, it is important to note that using data triangulation would have also helped reduce the negative impact of these limitations and improved the robustness of this study's findings as well.

Implications and Recommendations for Practice and Future Research

Findings of this investigation are accessible to researchers in different disciplines to build on, engineers working on the development of accommodating Web-based resources, clinicians who need to be informed about the challenges their patients face in everyday life, and policy makers who can create evidence-based policies that can together realize very positive outcomes for people with depression and anxiety. These findings also place the spotlight on the importance of considering difficulties associated with affective states when delivering enablement initiatives involving the use of technology. This is in contrast to the substantial degree of attention given to the needs of those with sensory impairments and physical disabilities. Researchers are also encouraged to adopt a more comprehensive view of accessibility that captures the complexity of the interaction between users and their environment.

The International Classification of Functioning, Disability, and Health's Model of Functioning and Disability provides a clear framework that can be used to describe and study the user experience of people with mental health conditions when using Web-based and other digital technologies. For example, the World Wide Web Consortium's Four Principles of Accessibility [57] focuses on instances where a Web-based resource is not understandable, perceivable, and operable but neglects the other important factors in the interaction between the user and the user's environment.

The high variability in user needs among people with depression and anxiety presents a unique challenge for accessibility professionals. Enablement efforts should be targeted at an individual level and no longer solely at a user group level. Meeting this challenge will call for new facilitation methods that rely on emerging technologies such as artificial intelligence to provide highly personalized experiences for each user with depression and anxiety and also other mental health conditions.

The adoption of stronger data protection policies (eg, [58]) will be of great benefit to people with depression and anxiety who worry about not having enough control over their personal data and having it be misused for privacy violations (eg, fear of unwanted direct contact) among other infractions.

Conclusions

People with depression and anxiety experience difficulties when using the Web that are related to sociocognitive deficits associated with their conditions. Participants in the comparison group did not experience most of these difficulties. MHPs are very aware of the difficulties that people with depression and anxiety are likely to experience when using the Web. Findings highlight several Web activities, services, and features that should be reviewed not only for people with depression and anxiety but also for people affected by other mental disorders and conditions that share similar symptomology. This investigation has contributed to a fuller understanding of these difficulties and provides guidance on what to address for people with depression and anxiety when using the Web. It also calls for novel approaches to aid in the removal and reduction of these difficulties using more carefully personalized experiences.

Acknowledgments

The research leading to these results has received funding from the People Programme (Marie Curie Actions) of the European Union's Seventh Framework Programme 2007-2013 under the Research Executive Agency grant agreement number 316795. The views expressed in this paper are those of the authors and do not necessarily represent the views or policies of the World Health Organization. The authors would like to thank professor David Baldwin, Dr Kaloyan Kamenov, and Dr Maria Cabello for their support during the data collection phase. The authors would also like to thank Simão Campos, Dr Kaloyan Kamenov, Dr Conal Twomey, and Mr Daley Holloway for their generous feedback during the final drafting stages.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Semistructured interview topic guide.
[PDF File (Adobe PDF File), 173 KB - jmir_v21i10e12514_app1.pdf]

Multimedia Appendix 2
Checklist for Reporting Results of Internet E-Surveys.
[PDF File (Adobe PDF File), 210 KB - jmir_v21i10e12514_app2.pdf]

Multimedia Appendix 3
Mental health practitioner expert Web-based survey questions.
[PDF File (Adobe PDF File), 218 KB - jmir_v21i10e12514_app3.pdf]

References

1. GBD 2016 Disease and Injury Incidence and Prevalence Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 328 diseases and injuries for 195 countries, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet 2017 Sep 16;390(10100):1211-1259 [FREE Full text] [doi: 10.1016/S0140-6736(17)32154-2] [Medline: 28919117]
2. Kalckreuth S, Trefflich F, Rummel-Kluge C. Mental health related Internet use among psychiatric patients: a cross-sectional analysis. BMC Psychiatry 2014 Dec 24;14:368 [FREE Full text] [doi: 10.1186/s12888-014-0368-7] [Medline: 25599722]
3. Chen J, Zhu S. Online information searches and help seeking for mental health problems in urban china. Adm Policy Ment Health 2016 Jul;43(4):535-545. [doi: 10.1007/s10488-015-0657-6] [Medline: 25981055]
4. Lau Y, Htun TP, Wong SN, Tam WS, Klainin-Yobas P. Therapist-supported internet-based cognitive behavior therapy for stress, anxiety, and depressive symptoms among postpartum women: a systematic review and meta-analysis. J Med Internet Res 2017 Apr 28;19(4):e138 [FREE Full text] [doi: 10.2196/jmir.6712] [Medline: 28455276]
5. Rooksby M, Elouafkaoui P, Humphris G, Clarkson J, Freeman R. Internet-assisted delivery of cognitive behavioural therapy (CBT) for childhood anxiety: systematic review and meta-analysis. J Anxiety Disord 2015 Jan;29:83-92. [doi: 10.1016/j.janxdis.2014.11.006] [Medline: 25527900]
6. Seabrook EM, Kern ML, Rickard NS. Social networking sites, depression, and anxiety: a systematic review. JMIR Ment Health 2016 Nov 23;3(4):e50 [FREE Full text] [doi: 10.2196/mental.5842] [Medline: 27881357]
7. Bernard R, Sabariego C, Cieza A. Barriers and facilitation measures related to people with mental disorders when using the web: a systematic review. J Med Internet Res 2016 Jun 9;18(6):e157 [FREE Full text] [doi: 10.2196/jmir.5442] [Medline: 27282115]
8. World Health Organization. The ICD-10 Classification Of Mental And Behavioural Disorders: Clinical Descriptions And Diagnostic Guidelines. Geneva: World Health Organization; 1992.
9. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders. Washington, DC: American Psychiatric Association; 2013.
10. Peter C, Beale R, Crane E, Axelrod L, Blyth G. Emotion in HCI - Designing for People. Fraunhofer IRB Verlag, Stuttgart: Joint Proceedings of the 2005, 2006, and 2007 Intl Workshops; 2008.
11. Feather JS, Howson M, Ritchie L, Carter PD, Parry DT, Koziol-McLain J. Evaluation methods for assessing users' psychological experiences of web-based psychosocial interventions: a systematic review. J Med Internet Res 2016 Jun 30;18(6):e181 [FREE Full text] [doi: 10.2196/jmir.5455] [Medline: 27363519]
12. Bernard R, Sabariego C, Baldwin D, Abou-Zahra S, Cieza A. BETTER-project: web accessibility for persons with mental disorders. In: Kurosu M, editor. Human-computer Interaction: Users And Contexts. Switzerland: Springer International Publishing; 2015:25-34.
13. Carter N, Bryant-Lukosius D, DiCenso A, Blythe J, Neville AJ. The use of triangulation in qualitative research. Oncol Nurs Forum 2014 Sep;41(5):545-547. [doi: 10.1188/14.ONF.545-547] [Medline: 25158659]
14. Palinkas LA, Horwitz SM, Green CA, Wisdom JP, Duan N, Hoagwood K. Purposeful sampling for qualitative data collection and analysis in mixed method implementation research. Adm Policy Ment Health 2015 Sep;42(5):533-544 [FREE Full text] [doi: 10.1007/s10488-013-0528-y] [Medline: 24193818]
15. Patton MQ. Qualitative Evaluation And Research Methods. Thousand Oaks, California: Sage Publications, Inc; 1990.
16. Hargittai E, Dobransky K. Old dogs, new clicks: digital inequality in skills and uses among older adults. Can J Commun 2017;42(2):195-212. [doi: 10.22230/cjc.2017v42n2a3176]
17. Hargittai E, Piper AM, Morris MR. From internet access to internet skills: digital inequality among older adults. Univ Access Inf Soc 2018:1-10. [doi: 10.1007/s10209-018-0617-5]
18. Hargittai E, Shaw A. Mind the skills gap: the role of Internet know-how and gender in differentiated contributions to Wikipedia. Inf Commun Soc 2015;18(4):424-442. [doi: 10.1080/1369118x.2014.957711]
19. Wong TY, Yuen KS, Li WO. A basic need theory approach to problematic internet use and the mediating effect of psychological distress. Front Psychol 2014;5:1562 [FREE Full text] [doi: 10.3389/fpsyg.2014.01562] [Medline: 25642201]

20. Guest G, Bunce A, Johnson L. How many interviews are enough? An experiment with data saturation and variability. Field Methods 2006;18(1):59-82. [doi: 10.1177/1525822X05279903]
21. Walker JL. The use of saturation in qualitative research. Can J Cardiovasc Nurs 2012;22(2):37-46. [Medline: 22803288]
22. Hargittai E, Hsieh YP. Succinct survey measures of web-use skills. Soc Sci Comput Rev 2011;30(1):95-107. [doi: 10.1177/0894439310397146]
23. Beck AT, Steer RA, Brown GK. BDI-II, Beck Depression Inventory: Manual. San Antonio, TX: Psychological Corporation; 1996.
24. Beck AT, Steer RA. Manual for the Beck Anxiety Inventory. San Antonio, TX: Psychological Corporation; 1993.
25. Kroenke K, Spitzer RL, Williams JB, Löwe B. An ultra-brief screening scale for anxiety and depression: the PHQ-4. Psychosomatics 2009;50(6):613-621. [doi: 10.1176/appi.psy.50.6.613] [Medline: 19996233]
26. Bernard HR. Social Research Methods: Qualitative And Quantitative Approaches. Thousand Oaks, California: Sage Publications Inc; 2012.
27. Ritchie J, Spencer L, Bryman A, Burgess R. Analysing qualitative data. In: Dey I, editor. Qualitative Data Analysis: A User Friendly Guide For Social Scientists. London: Routledge; 1993.
28. Rabiee F. Focus-group interview and data analysis. Proc Nutr Soc 2004 Nov;63(4):655-660. [doi: 10.1079/pns2004399] [Medline: 15831139]
29. Krueger RA, Casey MA. Focus Groups: A Practical Guide for Applied Research. Thousand Oaks, California: SAGE; 2000.
30. Etikan I, Bala K. Sampling and sampling methods. Biom Biostat Int J 2017;5(6):1-3. [doi: 10.15406/bbij.2017.05.00149]
31. Singh K. Quantitative Social Research Methods. Thousand Oaks, California: Sage Publication Inc; 2007.
32. Taylor PJ, Rietzschel J, Danquah A, Berry K. The role of attachment style, attachment to therapist, and working alliance in response to psychological therapy. Psychol Psychother 2015 Sep;88(3):240-253. [doi: 10.1111/papt.12045] [Medline: 25384637]
33. Eysenbach G. Improving the quality of Web surveys: the Checklist for Reporting Results of Internet E-Surveys (CHERRIES). J Med Internet Res 2004 Sep 29;6(3):e34 [FREE Full text] [doi: 10.2196/jmir.6.3.e34] [Medline: 15471760]
34. World Health Organization. International Classification Of Functioning, Disability And Health: ICF. Geneva: World Health Organization; 2001.
35. Cieza A, Anczewska M, Ayuso-Mateos JL, Baker M, Bickenbach J, Chatterji S, PARADISE Consortium. Understanding the impact of brain disorders: towards a 'horizontal epidemiology' of psychosocial difficulties and their determinants. PLoS One 2015;10(9):e0136271 [FREE Full text] [doi: 10.1371/journal.pone.0136271] [Medline: 26352911]
36. World Health Organization. The International Statistical Classification Of Diseases And Health Related Problems: ICD-10. Geneva: World Health Organization; 2004.
37. Sellen AJ, Murphy R, Shaw KL. How Knowledge Workers Use the Web. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2002 Presented at: CHI'02; April 20-25, 2002; Minneapolis, Minnesota, USA p. 227-234. [doi: 10.1145/503417.503418]
38. Kellar M, Watters C, Shepherd M. A Goal-based Classification of Web Information Tasks. In: Proceedings of the 69th Annual Meeting of the American Society for Information Science and Technology. 2006 Presented at: ASIST'06; November 3-8, 2006; Austin, US. [doi: 10.1002/meet.14504301121]
39. Watkins ER. Constructive and unconstructive repetitive thought. Psychol Bull 2008 Mar;134(2):163-206 [FREE Full text] [doi: 10.1037/0033-2909.134.2.163] [Medline: 18298268]
40. Segerstrom S, Tsao J, Alden L, Craske M. Worry and rumination: Repetitive thought as a concomitant and predictor of negative mood. Cogn Ther Res 2000;24(6):671-688. [doi: 10.1023/A:1005587311498]
41. Larijani HA, Garmaroudi G. Effectiveness of Resiliency Training in the Mental Health of the Red Crescent Health Care Teams. Revista Romaneasca pentru Educatie Multidimensionala 2018 Jul 02;10(1 Suppl):64-83. [doi: 10.18662/rrem/37]
42. Rutten BP, Hammels C, Geschwind N, Menne-Lothmann C, Pishva E, Schruers K, et al. Resilience in mental health: linking psychological and neurobiological perspectives. Acta Psychiatr Scand 2013 Jul;128(1):3-20 [FREE Full text] [doi: 10.1111/acps.12095] [Medline: 23488807]
43. Papageorgiou C, Wells A. Depressive Rumination: Nature, Theory and Treatment. Hoboken, NJ: John Wiley & Sons; 2004.
44. Chou HG, Edge N. 'They are happier and having better lives than I am': the impact of using Facebook on perceptions of others' lives. Cyberpsychol Behav Soc Netw 2012 Feb;15(2):117-121. [doi: 10.1089/cyber.2011.0324] [Medline: 22165917]
45. Lee SY. How do people compare themselves with others on social network sites?: The case of Facebook. Computers in Human Behavior 2014;32:253-260. [doi: 10.1016/j.chb.2013.12.009]
46. Dittmar H, Howard S. Thin-ideal internalization and social comparison tendency as moderators of media models' impact on women's body-focused anxiety. J Soc Clin Psychol 2004 Dec;23(6):768-791. [doi: 10.1521/jscp.23.6.768.54799]
47. Chen W, Lee K. Sharing, liking, commenting, and distressed? The pathway between Facebook interaction and psychological distress. Cyberpsychol Behav Soc Netw 2013 Oct;16(10):728-734. [doi: 10.1089/cyber.2012.0272] [Medline: 23745614]
48. Rosenthal SR, Buka SL, Marshall BDL, Carey KB, Clark MA. Negative experiences on Facebook and depressive symptoms among young adults. J Adolesc Health 2016 Nov;59(5):510-516. [doi: 10.1016/j.jadohealth.2016.06.023] [Medline: 27546886]

49. Valkenburg PM, Peter J. Social consequences of the internet for adolescents. Curr Dir Psychol Sci 2009;18(1):1-5. [doi: 10.1111/j.1467-8721.2009.01595.x]
50. Hampton K, Rainie L, Lu W, Shin I, Purcell K. Pew Research Center. 2015. Social Media and the Cost of Caring URL: https://www.pewinternet.org/2015/01/15/social-media-and-stress/ [accessed 2019-10-08]
51. Liu C, Ang RP, Lwin MO. Cognitive, personality, and social factors associated with adolescents' online personal information disclosure. J Adolesc 2013 Aug;36(4):629-638. [doi: 10.1016/j.adolescence.2013.03.016] [Medline: 23849657]
52. Vannucci A, Flannery KM, Ohannessian CM. Social media use and anxiety in emerging adults. J Affect Disord 2017 Jan 1;207:163-166. [doi: 10.1016/j.jad.2016.08.040] [Medline: 27723539]
53. Andreassen CS, Pallesen S. Social network site addiction - an overview. Curr Pharm Des 2014;20(25):4053-4061. [doi: 10.2174/13816128113199990616] [Medline: 24001298]
54. Zhang C, Brook JS, Leukefeld CG, Brook DW. Associations between compulsive buying and substance dependence/abuse, major depressive episodes, and generalized anxiety disorder among men and women. J Addict Dis 2016;35(4):298-304 [FREE Full text] [doi: 10.1080/10550887.2016.1177809] [Medline: 27215919]
55. Gallagher CE, Watt MC, Weaver AD, Murphy KA. 'I fear, therefore, I shop!' exploring anxiety sensitivity in relation to compulsive buying. Pers Individ Dif 2017;104:37-42. [doi: 10.1016/j.paid.2016.07.023]
56. Swar B, Hameed T, Reychav I. Information overload, psychological ill-being, and behavioral intention to continue online healthcare information search. Comput Human Behav 2017;70:416-425. [doi: 10.1016/j.chb.2016.12.068]
57. World Wide Web Consortium. 2016. Introduction to Understanding WCAG 2.0 URL: https://www.w3.org/TR/UNDERSTANDING-WCAG20/intro.html [accessed 2018-10-14] [WebCite Cache ID 73Ady0LpD]
58. Council of the European Union, European Parliament. Publications Office of the EU. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC URL: https://publications.europa.eu/en/publication-detail/-/publication/3e485e15-11bd-11e6-ba9a-01aa75ed71a1/language-en [accessed 2019-10-08]

Abbreviations

MHP: mental health practitioner
DA: depression and anxiety

Edited by G Eysenbach; submitted 28.10.18; peer-reviewed by A Dodd, J Mueller, SJ Kim, Y Lau; comments to author 21.03.19; revised version received 31.07.19; accepted 31.08.19; published 31.10.19.

Please cite as:
Bernard R, Sabariego C, Cieza A
Difficulties Encountered by People With Depression and Anxiety on the Web: Qualitative Study and Web-Based Expert Survey
J Med Internet Res 2019;21(10):e12514
URL: http://www.jmir.org/2019/10/e12514/
doi:10.2196/12514
PMID:31674915

Original Paper

Patient Engagement in Medical Research Among Older Adults: Analysis of the Health Information National Trends Survey

Lynette Hammond Gerido[1], MPH; Xiang Tang[2], MS; Brittany Ernst[3]; Aisha Langford[4], MPH, PhD; Zhe He[1], MS, PhD

[1]School of Information, Florida State University, Tallahassee, FL, United States

[2]Department of Statistics, Florida State University, Tallahassee, FL, United States

[3]College of Human Sciences, Florida State University, Tallahassee, FL, United States

[4]Department of Population Health, School of Medicine, New York University, New York, NY, United States

Corresponding Author:
Zhe He, MS, PhD
School of Information
Florida State University
142 Collegiate Loop
Tallahassee, FL, 32306
United States
Phone: 1 8506445775
Email: zhe@fsu.edu

Abstract

Background: By 2035, it is expected that older adults (aged 65 years and older) will outnumber children and will represent 78 million people in the US population. As the aging population continues to grow, it is critical to reduce disparities in their representation in medical research.

Objective: This study aimed to describe sociodemographic characteristics and health and information behaviors as factors that influence US adults' interest in engaging in medical research, beyond participation as study subjects.

Methods: Nationally representative cross-sectional data from the 2014 Health Information National Trends Survey (N=3677) were analyzed. Descriptive statistics and weighted multivariable logistic regression analyses were performed to assess predictors of one's interest in patient engagement in medical research. The independent variables included age, general health, income, race and ethnicity, education level, insurance status, marital status, and health information behaviors.

Results: We examined the association between the independent variables and patient interest in engaging in medical research (PTEngage_Interested). Patient interest in engaging in medical research has a statistically significant association with age (adjusted $P<.01$). Younger adults (aged 18-34 years), lower middle-aged adults (aged 35-49 years), and higher middle-aged adults (aged 50-64 years) indicated interest at relatively the same frequency (29.08%, 29.56%, and 25.12%, respectively), but older adults (aged ≥65 years) expressed less interest (17.10%) than the other age groups. After the multivariate model was run, older adults (odds ratio 0.738, 95% CI 0.500-1.088) were found to be significantly less likely to be interested in engaging in medical research than adults aged 50 to 64 years. Regardless of age, the strongest correlation was found between interest in engaging in medical research and actively looking for health information ($P<.001$). Respondents who did not seek health information were significantly less likely than those who did seek health information to be interested in engaging in medical research.

Conclusions: Patients' interest in engaging in medical research vary by age and information-seeking behaviors. As the aging population continues to grow, it is critical to reduce disparities in their representation in medical research. Interest in participatory research methods may reflect an opportunity for consumer health informatics technologies to improve the representation of older adults in future medical research.

(J Med Internet Res 2019;21(10):e15035) doi:10.2196/15035

KEYWORDS

aging; health care disparities; patient participation; medical informatics

Introduction

In the United States, the older population is growing as life expectancy increases, and the *baby boomers*, those born in post-World War II America, reach age 65 years and beyond [1]. By 2035, it is expected that older people will outnumber children and will represent 78 million people in the US population [2]. Generally, increasing life expectancy is considered a positive human development; however, growing older is inherently associated with biological and cognitive degeneration [3]. Deteriorating physical health among older adults is most likely because of an increasing prevalence of chronic conditions such as hypertension, rheumatoid arthritis, heart failure, diabetes, lung disease, cancer, and mood and anxiety disorders [1,4,5].

In the United States, 80% of people aged 65 years and older suffer from multiple chronic conditions [6]. Compared with their younger counterparts, older people have increased rates of comorbidities and complications [5,7]. Although older patients are increasing in number, research has shown that they have been underrepresented in medical research such as clinical trials [8-10]. As such, a clear health disparity exists because older populations are unable to benefit from innovative technologies and treatments that may improve health outcomes, improve quality of life, or reduce their overall disease burden [11].

Interest in enhancing patient engagement in medical research is growing, and patients are increasingly playing the role of active partners invested in better health outcomes [12]. Researchers have been working to tailor research agendas to reduce disparities by developing systems and processes to directly involve patients and communities in research [13-15]. Several studies have identified patient- and community-level barriers to participation, noting a lack of understanding of the benefits of clinical trial participation [16,17]. Therefore, patient engagement in research proactively employs collaborative approaches to inquiry or investigation [18]. When patients are engaged in medical research, the goal is for the patients to clearly understand their role in the research process [19] and to be continuously updated about advances resulting from medical research [20]. Patient engagement has been recognized in the literature for its potential benefits such as improvement in the credibility of results related to higher rates of participation, direct applicability of results, improved translation of results into practice, and advances through the ethical focus on *democratization* of research [21-23]. In addition to participating in clinical trials, patients may engage in medical research by performing very specific roles such as serving on a community advisory board or spokesmanship such as being the *public face* of the project [21]. Informatics researchers have applied patient engagement practices to explore community technology practices, develop technology interventions, and use technology to understand community problems [24].

As more researchers implement participatory research design strategies, informatics professionals may further examine the level of engagement of underserved populations in medical research in comparison with the diffusion of technologies such as direct patient engagement networks as tools for information sharing and participation. In 2016, the US Congress passed the 21st Century Cures Act [25], which authorized US $1.8 billion in funding for the Cancer Moonshot Initiative to accelerate advances in cancer research [26]. The Cancer Moonshot Initiative relies on the recommendations from a Blue Ribbon Panel (BRP) of scientific experts to advise the National Cancer Board on actions the broader community views most able to accelerate research [27]. Reducing cancer disparities across a range of research areas using a cross-cutting theme of patient engagement has been at the forefront of the BRP recommendations [28]. The BRP identified direct patient engagement in cancer research including access to a network of information and tools for data sharing as a promising approach to advancing research [29]. In 2010, the Patient-Centered Outcomes Research Institute (PCORI) was established to fund comparative clinical effectiveness research (CER) with an emphasis on answering questions important to patients by recruiting a large and diverse patient population, thereby assisting patients in making informed health decisions [30]. The PCORI developed a National Patient-Centered Clinical Research Network (PCORnet) to improve the nation's capacity to conduct CER. This national resource is referred to as a *network of networks* and includes 18 patient-powered research networks [31].

When patients have easily available, accurate, and timely information and use it to make informed choices, they are empowered and experience improved health outcomes [32,33]. In participatory research designs, patients may be both producers and consumers of information with opportunities to build local, grassroots action networks for information dissemination [34]. Despite the intuitive appeal of patient engagement strategies, their efficacy for reducing disparities remains inconclusive [30]. To date, there is little scientific evidence available addressing how interest in engaging in medical research among older adults compares with that of younger generations. Although it is certain that there will be a profound impact on the use of health information and technology by the greater number of older people in the population, an important question for researchers is whether this impact might be larger or smaller because health behaviors and characteristics of older adults are different from those of their younger counterparts. Thus, this study aimed to (1) describe current levels of interest in engaging in medical research in the United States and (2) identify patient-level sociodemographic, behavioral, and health information–seeking characteristics associated with interest in engagement in medical research. Few studies have evaluated patient engagement in medical research using a nationally representative sample, which may help assess the impact of a broadly applicable framework [35]. Devising a baseline of patient engagement in medical research and understanding the sociodemographic characteristics associated with engagement of older people can improve future efforts to increase participation of older adults in medical research and thereby reduce disparities in health outcomes.

Methods

Data Collection

We derived data for this analysis from the National Cancer Institute's Health Information National Trends Survey (HINTS). HINTS is a nationally representative cross-sectional survey administered biennially to adults aged 18 years and older in the United States to monitor the evolution of health information and communication. Each version of HINTS includes slightly different survey questions. We used HINTS 4 Cycle 4 data collected between August and November 2014 via self-administered mailed questionnaires. The survey response rate was 34.4%. Additional HINTS 4 methodology details have been described elsewhere [36].

The deidentified HINTS 4 Cycle 4 dataset is publicly available [37]. At the time of the analysis, HINTS 4 Cycle 4 was the most recent iteration for which responses were available for the following questions:

> *More and more, people are getting involved in research in new ways beyond being a research subject. They are partnering with medical researchers to help decide what research is done and how it is done. For example, people can suggest important topics to study or how to report results to the public. This is sometimes called "patient engagement" in research.*
>
> *1. Have you ever heard about "patient engagement" in medical research?*
>
> *2. Have you ever engaged in medical research in this way?*
>
> *3. Would you ever be interested in engaging in research in this way?*

This population sample included 3677 respondents. The exclusion criteria for this study were based on responses to age. Respondents were excluded if they had a missing response or a response error (n=182).

Measures

Dependent Variables

The study had 3 key outcome variables related to patient engagement in medical research, which were assessed as awareness ("Have you ever heard about 'patient engagement' in medical research?" [PTEngage_Heardof]), past experience ("Have you ever engaged in medical research in this way?" [PTEngage_EverEngaged]), and current interest ("Would you ever be interested in engaging in research in this way?" [PTEngage_Interested]). Answer choices for these variables are *yes*, *no*, or *not sure*.

Independent Variables

The following variables were tested for independence and correlation. They are grouped as sociodemographics, health-related variables, and health information behavior variables.

Sociodemographics

We examined selected self-reported sociodemographic characteristics including gender (male / female), age (18-34 / 35-49 / 50-64 / ≥65 years), income (<US $19,999 / US $20,000-US $34,999 / US $35,000-US $49,999 / US $50,000-US $74,999 / ≥US $75,000), race and ethnicity (non-Hispanic white / non-Hispanic black / Hispanic / Asian / Other), education (less than high school / high school graduate / some college / college graduate and more), occupation (employed / unemployed / homemaker / student / retired / disabled), marital status (married/living as married / divorced / widowed / separated / single), rurality, and active duty military service. Rurality (yes/no) was determined by HINTS variable *RUC2013*, which uses the 2013 USDA rural / urban designation assigned to the respondent's mailing address. Active duty military service was recoded to a bivariate (yes / no) using the responses for the following question: "Have you ever served on active duty in the US Armed Forces, military Reserves, or National Guard?" A variable describing emotional support and self-efficacy was included as well. Social support (yes / no) acknowledges having someone with whom to talk about problems and to help with decision making.

Health-Related Variables

Potentially important clinical characteristics and health behaviors such as cancer history, general health, health insurance, regular provider, and most recent checkup were included. Respondents were asked about their health behaviors such as whether or not they have a regular health care provider and how long it has been since their last routine checkup. These questions allowed responses of *yes*, *no*, or *not sure*. In addition, respondents were asked to rate their confidence in their own ability to take good care of their health using a 5-point Likert scale with 1 being *completely confident* and 5 being *not confident at all*.

Health Information Behavior Variables

Respondents were also asked whether they had ever looked for health information from any source, which was a binary *yes* or *no* response. They were asked to qualify their health information seeking by indicating for whom the information was sought (ie, *self*, *someone else*, or both) and what sources (eg, books, family, internet, and library) were used. In addition, the respondents were asked what technologies they used to exchange digital health information with their provider. The options (ie, email, text message, app on mobile device, video conference, or fax) were presented with binary *yes* or *no* responses. We used the derived MedINfo_Cat variable to categorize the responses (ie, email only, text message only, app on smartphone only, video or social media only, fax only, none, or multiple technologies).

We included the responses to questions related to access to electronic medical records and personal health records. The binary (yes/no) question "As far as you know do any of your doctors or health care providers maintain your medical information in a computerized system?" was included to determine provider users of electronic medical records. In addition, we included responses indicating the importance (very important, somewhat important, and not at all important) for

the statement, "You should be able to get to your own medical information electronically." Finally, we included responses indicating confidence (very confident, somewhat confident, and not confident) to the question "How confident are you that you have some say in who is allowed to collect, use, and share your medical information?"

Statistical Analysis

We conducted a complete case analysis and used SAS software (SAS Institute Inc), version 9.4 to perform all statistical analyses. First, we evaluated the frequencies of sociodemographics and health information behavior variables (see Multimedia Appendix 1).

Then we used the chi-square test to examine associations between sociodemographic and health information behavior variables with each outcome of patient engagement: awareness (PTEngage_HeardOf), past experience (PTEngage_EverEngaged), and current interest (PTEngage_Interested; see Multimedia Appendix 1). We found the awareness (PTEngage_HeardOf) and past experience (PTEngage_EverEngaged) variables to have little statistically significant association with age (adjusted P=.22 and P=.88, respectively). Therefore, we focused on examining the association between independent variables with a single outcome of a single dependent variable for patient engagement and current interest (PTEngage_Interested), which had an adjusted P=.001 and indicates a statistically significant association with age.

We conducted multivariate logistic regression analyses to examine whether the health information behaviors variables that were significantly associated with patient engagement remained significant when controlling for certain sociodemographic factors. Weighted multivariate logistic regression analyses were used to obtain an odds ratio (OR) and 95% CI (see Multimedia Appendix 2). A full model, which included all the variables of interest, was developed, and backward elimination was used to identify covariates that were significantly correlated and influenced the regression estimates. Then we revised the model to include only primary independent variables that were significant in the initial model. A statistical significance criterion of P<.05 was used for all analyses. Owing to missing data, sample sizes for these multivariate analyses ranged from 2463 to 3495.

To account for the HINTS sampling design and calculate nationally representative estimates, we applied SAS survey procedures incorporating the jackknife variance estimation technique and HINTS-supplied survey weights. This study was granted an expedited ethical approval by the Florida State University Institutional Review Board.

Results

Sociodemographic Characteristics

We defined 4 age groups as follows: younger adults (aged 18-34 years), lower middle-aged adults (aged 35-49 years), higher middle-aged adults (aged 50-64 years), and older adults (aged ≥65 years). Each of the sociodemographic characteristics differed significantly with respect to our age groups (Table 1). The full results of the Chi-square tests are provided in Multimedia Appendix 1.

Most study respondents were aged between 50 and 64 years (higher middle-aged), married, college educated, white, lived in urban areas, and homeowners and had an annual household income over US $50,000. Older adults tended to be white (80.00%) and less diverse than the other age groups. Younger adults (aged 18-34 years) tended to have higher educational attainment (50.27%, college or higher) than older adults (27.13%, college or higher). Lower middle-aged, higher middle-aged, and older adults had larger married representation (66.95%, 67.82%, and 56.24%, respectively) than younger adults who were predominately single and never married (28.79% married and 67.39% single). Older adults had the highest representation of widowers (24.90%). Among younger adults, a vast majority reported having emotional support (92.69%) or friends and family to talk with about health (89.17%). Similarly, among older adults most reported having emotional support (88.50%) and having friends or family to talk with about health (90.72%). The older adults have served in the military for active duty more than any other age group (23.87%).

Table 1. Weighted percentage of individual characteristics by age group.

Variable	Younger adults (aged 18-34 years)	Lower middle-aged adults (aged 35-49 years)	Higher middle-aged adults (aged 50-64 years)	Older adults (aged ≥65 years)	All
Race and ethnicity[a] (n=3229), n (weighted %)					
Non-Hispanic white	225 (58.19)	333 (62.38)	701 (73.73)	651 (80.00)	1940 (34.05)
Non-Hispanic black	61 (13.26)	144 (12.53)	215 (10.43)	103 (6.67)	523 (25.30)
Hispanic	89 (16.76)	176 (19.88)	148 (10.73)	118 (10.50)	531 (23.98)
Non-Hispanic Asian	27 (7.69)	35 (3.70)	37 (3.69)	23 (2.54)	122 (22.15)
Other	25 (4.09)	29 (1.51)	44 (1.41)	15 (0.39)	113 (15.59)
Gender[b] (n=3460), n (weighted %)					
Male	141 (50.00)	265 (48.87)	495 (48.03)	468 (43.40)	1369 (48.05)
Female	323 (50.00)	472 (51.13)	714 (51.97)	582 (56.60)	2091 (51.95)
Education[a] (n=3470), n (weighted %)					
Less than high school	19 (5.05)	67 (11.87)	97 (12.93)	122 (21.04)	305 (11.62)
12 years or completed high school	58 (12.50)	105 (19.27)	243 (19.34)	248 (24.32)	654 (18.08)
Some college	143 (32.17)	192 (26.92)	425 (32.43)	311 (27.51)	1071 (30.02)
College graduate or higher	245 (50.27)	376 (41.92)	450 (35.30)	369 (27.13)	1440 (40.28)
Income range[a] (n=3166), n (weighted %)					
<US $19,999	99 (22.53)	120 (13.64)	278 (17.31)	235 (25.33)	735 (19.26)
US $20,000-US $34,999	67 (12.84)	73 (8.58)	143 (11.62)	189 (21.23)	472 (12.72)
US $35,000-US $49,999	72 (14.59)	103 (16.15)	150 (12.56)	142 (16.04)	467 (14.74)
US $50,000-US $74,999	94 (17.90)	123 (16.47)	181 (18.01)	139 (16.05)	537 (17.24)
≥US $75,000	117 (32.13)	287 (45.16)	362 (40.50)	192 (21.35)	958 (36.03)
Have health insurance[a] (n=3446), n (weighted %)					
Yes	369 (82.62)	626 (86.78)	1032 (86.57)	1014 (98.20)	3041 (87.41)
Have a regular health care provider[a] (n=3439), n (weighted %)					
Yes	227 (47.01)	449 (63.4)	886 (74.42)	855 (81.93)	2417 (64.20)
How many times did you access your own personal health information online through a secure website or app in the last 12 months?[c] (n=3438), n (weighted %)					
None	312 (68.34)	498 (69.64)	871 (74.33)	844 (81.32)	2525 (74.45)
1-2 times	72 (16.57)	118 (14.83)	144 (10.58)	107 (10.06)	441 (13.47)
3-5 times	42 (6.70)	72 (8.96)	88 (7.59)	52 (4.86)	254 (7.21)
6-9 times	16 (3.09)	24 (3.94)	48 (3.33)	23 (1.87)	111 (3.16)
≥10 times	21 (5.30)	20 (2.64)	45 (4.16)	21 (1.89)	107 (3.71)
Seek health information[c] (n=3463), n (weighted %)					
Yes	385 (78.64)	634 (84.53)	994 (82.64)	825 (77.60)	2838 (81.04)
Interested in medical research[b] (n=3368), n (weighted %)					
Yes	148 (29.08)	216 (29.56)	329 (25.17)	191 (17.10)	884 (23.15)

[a]Statistically significant, $P<.001$.

[b]Statistically significant, $P<.01$.

[c]Statistically significant, $P<.05$.

Health Information Behaviors

Nearly all older adults (98.20%) reported having health insurance. Similarly, 85.41% of older adults indicated having had a checkup in the past year and 81.93% confirmed having a regular health care provider. With respect to having confidence in their own ability to take care of their health, all age groups selected “very confident” more than any other option (younger adults, 44.25%; lower middle-aged, 47.07%; higher middle-aged, 45.58%; and older adults, 48.41%). Health information–seeking activities were present among all groups. Younger and older adults least frequently sought health information (78.64% and 77.60%, respectively), whereas lower middle-aged and higher middle-aged adults tended to report seeking health information slightly more (84.53% and 82.64%, respectively). Among those who sought health information, all age groups were more likely to report using the internet more than any other method, but a smaller proportion of older adults use the internet to seek health information (46.93%). In addition to the internet, older adults seek health information from their doctor (26.16%), which is a greater proportion than any other age group’s second most popular information channel. Although it was not found to be a statistically significant independent variable, a majority of all age groups used no technology to exchange medical information with health care professionals. Yet, every age group placed a high level of importance on an individual ability to get personal medical information electronically.

Patient Engagement

Regarding interest in engaging in medical research, younger, lower middle-aged, and higher middle-aged adults indicated interest at relatively the same frequency (29.08%, 29.56%, and 25.17%, respectively), but older adults expressed slightly less interest (17.10%) than the other age groups. A multivariate model was run using interest in engagement in medical research as the dependent variable and the remaining variables as independent variables (see Multimedia Appendix 1). Variables that were significant in the initial model were included in the final model (see Multimedia Appendix 2).

Compared with higher middle-aged adults, older adults (OR 0.738, 95% CI 0.500-1.088) were significantly less likely to be interested in engaging in medical research (Table 2). Many sociodemographic characteristics were not significantly associated with interest in engaging in research, but a few characteristics related to health behaviors and access to health data were found to be statistically significant. When controlling for all other variables, not having a regular health care provider reduces the odds of interest in engaging in medical research (OR 0.643, 95% CI 0.419-0.986). In addition, respondents who accessed their personal health information 6 to 9 times through a secure website or app in the last 12 months were 3 times as likely to be interested in engaging in medical research than those who had not accessed their personal health information. Finally, the strongest correlation was found between interest in engaging in medical research and actively looking for health information ($P<.001$). Respondents who did not seek health information were significantly less likely than those who did seek health information to be interested in engaging in medical research.

Table 2. Interest in medical research by sociodemographic characteristics, health behaviors, and information-seeking correlates.

Variable	Odds ratio	95% Wald confidence limits	*P* value
Age group (years)			
Younger, 18-34	1.260	0.729-2.179	.37
Low middle, 35-49	1.315	0.886-1.951	.12
High middle, 50-64 (reference)	1	__[a]	—
Older, ≥65	0.738	0.500-1.088	.046[b]
Have a regular health care provider			
Yes (reference)	1	—	—
No	0.643	0.419-0.986	.04[b]
How many times did you access your own personal health information online through a secure website or app in the last 12 months?			
None (reference)	1	—	—
1-2 times	1.161	0.689-1.957	.08
3-5 times	1.388	0.773-2.492	.44
6-9 times	3.064	1.523-6.167	.02[b]
≥10 times	2.586	1.175-5.692	.18
Seek health information			
Yes (reference)	1	—	—
No	0.253	0.134-0.476	<.001[c]

[a]Not applicable for references.

[b]Statistically significant, $P<.05$.

[c]Statistically significant, $P<.001$.

Discussion

Principal Findings

This study reports the prevalence of patients' interest in engaging in medical research using data collected from the 2014 HINTS. Using these nationally representative data, we were able to explicate relationships between specific sociodemographic characteristics, health behaviors, and information-seeking activities. The key finding from our analyses was that the association of age with interest in engaging in medical research remained significant after adjustment for potential confounders. In addition, we found that having a regular health care provider, accessing your personal health information 6 to 9 times per year, and seeking health information increased the odds of being interested in engaging in medical research.

Older Adults

Levy and Sidel [38] described social injustice conceptually as the denial of economic, sociocultural, political, civil, or human rights to individuals based on the perception of their inferiority by those with more power or influence. Operating under this definition of social injustice, it is the charge of our society to implement policies and actions to counter injustice [39]. We must ensure conditions under which people can be healthy. Therefore, it is the societal duty to strongly recommend greater investments in aging research and translation of study results into safe, affordable, and universally available applied technologies and treatments [3]. Previous research showed that older adults are often directly (with age criterion) excluded from participation in clinical trials [40]. The result of this study shows that older adults (aged ≥65 years) are less interested in engaging in medical research than middle-aged adults. Opportunities for the engagement in medical research by older populations may serve as a means to advocate and prioritize the needs of older adults to reduce disparities. Researchers may use informatics tools to assess priorities of information tracking systems, institution infrastructure, research infrastructure, navigator and personnel programs, and community partnerships and patient advocates relevant to older adults to increase access to education, screening, and research participation opportunities [41].

Regular Health Care Provider

In this study, we found that having a regular health care provider significantly increases the likelihood of interest in engaging in medical research. A number of factors could be in play. As many patients rely on their health care providers as their first choice for health information, providers may act as gatekeepers for information about medical research. Physicians may also be researchers and enrich the quality of both services and research studies [42]. However, it should be noted that the strongest arguments against gatekeeping center on the patient's lack of freedom, lack of choice, and the erosion of patient-doctor trust that springs from the doctor's prerogative to decide on any referral [43]. Another concept could be that patients with a

regular provider are actively treating a disease, and because of this, the patient is more familiar and interested in medical research. For example, for cancer patients, diagnosis represents a communications and information flow between providers and patients to support informed decision making [44]. Diagnosis may include consultations and counseling to determine the best treatments and opportunities to participate in research [44]. Despite the prevalence of the internet, social media, and smartphone apps, patients tend to rely most on physicians as a source of information on cancer and medical research [45].

Health Information Behavior

Acquiring and making sense of health information is vital to patients making important decisions for their own health and the health of their families [46]. Medicine is an information-intensive enterprise [47]. To make informed choices and navigate within a complex health care system, consumers must have easily available, accurate, and timely information, and they must use it [33]. Moreover, 2 key health information covariates, accessing personal health information 6 to 9 times a year and seeking health information, appear to be correlated with interest in engaging in medical research. Both may be related to patient activation, which refers to empowering patients to play an active role in health care [12,48]. When patients are provided with access to their health information, they tend to have higher levels of satisfaction with their providers, increased understanding of their care, more engagement in health improving behaviors, and improved health outcomes [49]. Consumer health information technologies may help improve provider-to-patient communication, health monitoring, and information access to support self-care [50]. Although older adults are traditionally *late adopters* of technology, many use the internet and mobile devices to seek out health information [51]. Community-engaged health informatics, which combines concepts and methods from biomedical informatics, community-based public health approaches, and community informatics, may present opportunities for informatics advancements in patient engagement in medical research [52]. In addition, with advances in informatics methods such as Generalizability Index for Study Trait [53] and tools such as Visual Analysis Tool of Clinical Study Target Population [54], medical researchers are discovering innovative ways to reduce disparities through improvement in population representativeness of their studies.

Limitations

This study is limited in 2 ways. First, HINTS does not include questions directly qualifying the respondents' interest in engaging in medical research. Future research should address this gap to better understand specifically what activities related to patient engagement in medical research were of interest, how they learned about opportunities for patients to engage in medical research, and whether they face any barriers to engaging. Second, as HINTS is a cross-sectional survey, it is not possible to infer causal relationships between variables. Despite these limitations, these findings present an opportunity to further explore differences among age groups, to better understand if specific behaviors related to patient engagement in medical research are of interest to older adults, and to identify how older adults learn about opportunities to engage in medical research. Future research may seek to describe additional factors, such as psychosocial influences, spatial or geographic trends, or diffusion of innovative consumer health information technologies. Such inquiry would benefit from a survey tool specifically designed to assess the level of patient interest in engaging in medical research.

Conclusions

The results of this study demonstrate that disparities exist among older adults with respect to interest in engaging in medical research. Advances in consumer health informatics within existing health systems and research agendas show promise. Future studies should focus on identifying optimal information systems for engaging patients and rigorously examining the impact of these tools for patient engagement in medical research. Specifically, there is a clear need for both methodological and practical research on patient engagement in medical research that translates to improved health outcomes and reduced disparities.

Acknowledgments

This project was partially supported by the National Institute on Aging of the National Institutes of Health (NIH) under Award Number R21AG061431 and the University of Florida Clinical and Translational Science Institute, which is supported in part by the NIH National Center for Advancing Translational Sciences under award number UL1TR001427. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Weighted percentage of individual characteristics by age group.
[PDF File (Adobe PDF File), 60 KB - jmir_v21i10e15035_app1.pdf]

Multimedia Appendix 2
Interest in medical research by sociodemographic characteristics, health behaviors, and information-seeking correlates.

[PDF File (Adobe PDF File), 51 KB - jmir_v21i10e15035_app2.pdf]

References

1. Cleary M, Sayers J, Bramble M, Jackson D, Lopez V. Overview of substance use and mental health among the 'baby boomers' generation. Issues Ment Health Nurs 2017 Jan;38(1):61-65. [doi: 10.1080/01612840.2016.1243177] [Medline: 27936334]
2. The United States Census Bureau. Older People Projected to Outnumber Children for First Time in US History URL: https://www.census.gov/newsroom/press-releases/2018/cb18-41-population-projections.html [accessed 2019-02-05]
3. Jin K, Simpkins J, Ji X, Leis M, Stambler I. The critical need to promote research of aging and aging-related diseases to improve health and longevity of the elderly population. Aging Dis 2015 Feb;6(1):1-5 [FREE Full text] [doi: 10.14336/AD.2014.1210] [Medline: 25657847]
4. King DE, Matheson E, Chirina S, Shankar A, Broman-Fulks J. The status of baby boomers' health in the United States: the healthiest generation? JAMA Intern Med 2013 Mar 11;173(5):385-386. [doi: 10.1001/jamainternmed.2013.2006] [Medline: 23381505]
5. He Z, Bian J, Carretta H, Lee J, Hogan W, Shenkman E, et al. Prevalence of multiple chronic conditions among older adults in Florida and the United States: comparative analysis of the OneFlorida data trust and national inpatient sample. J Med Internet Res 2018 Apr 12;20(4):e137 [FREE Full text] [doi: 10.2196/jmir.8961] [Medline: 29650502]
6. Gerteis J, Izrael D, Deitz D, LeRoy L, Ricciardi R, Miller T. The Agency for Healthcare Research and Quality. Multiple Chronic Conditions Chartbook URL: https://www.ahrq.gov/sites/default/files/wysiwyg/professionals/prevention-chronic-care/decision/mcc/mccchartbook.pdf [accessed 2019-01-19]
7. Janssen-Heijnen M, Houterman S, Lemmens V, Louwman M, Maas H, Coebergh J. Prognostic impact of increasing age and co-morbidity in cancer patients: a population-based approach. Crit Rev Oncol Hematol 2005 Sep;55(3):231-240. [doi: 10.1016/j.critrevonc.2005.04.008] [Medline: 15979890]
8. He Z, Charness N, Bian J, Hogan W. Assessing the comorbidity gap between clinical studies and prevalence in elderly patient populations. IEEE EMBS Int Conf Biomed Health Inform 2016 Feb;2016:136-139 [FREE Full text] [doi: 10.1109/BHI.2016.7455853] [Medline: 27738664]
9. Kennedy-Martin T, Curtis S, Faries D, Robinson S, Johnston J. A literature review on the representativeness of randomized controlled trial samples and implications for the external validity of trial results. Trials 2015 Nov 3;16:495 [FREE Full text] [doi: 10.1186/s13063-015-1023-4] [Medline: 26530985]
10. Lewis JH, Kilgore ML, Goldman DP, Trimble EL, Kaplan R, Montello MJ, et al. Participation of patients 65 years of age or older in cancer clinical trials. J Clin Oncol 2003 Apr 1;21(7):1383-1389. [doi: 10.1200/JCO.2003.08.010] [Medline: 12663731]
11. Guadagnolo B, Petereit D, Helbig P, Koop D, Kussman P, Dunn E, et al. Involving American Indians and medically underserved rural populations in cancer clinical trials. Clin Trials 2009 Dec;6(6):610-617 [FREE Full text] [doi: 10.1177/1740774509348526] [Medline: 19933720]
12. Menichetti J, Libreri C, Lozza E, Graffigna G. Giving patients a starring role in their own care: a bibliometric analysis of the on-going literature debate. Health Expect 2016 Jun;19(3):516-526 [FREE Full text] [doi: 10.1111/hex.12299] [Medline: 25369557]
13. Brett J, Staniszewska S, Mockford C, Herron-Marx S, Hughes J, Tysall C, et al. Mapping the impact of patient and public involvement on health and social care research: a systematic review. Health Expect 2014 Oct;17(5):637-650 [FREE Full text] [doi: 10.1111/j.1369-7625.2012.00795.x] [Medline: 22809132]
14. Michaels M, Weiss E, Guidry J, Blakeney N, Swords L, Gibbs B, et al. 'The promise of community-based advocacy and education efforts for increasing cancer clinical trials accrual'. J Cancer Educ 2012 Mar;27(1):67-74. [doi: 10.1007/s13187-011-0271-6] [Medline: 21938600]
15. Carter A, Nguyen C. A comparison of cancer burden and research spending reveals discrepancies in the distribution of research funding. BMC Public Health 2012 Jul 17;12:526. [doi: 10.1186/1471-2458-12-526] [Medline: 22800364]
16. Featherstone K, Donovan J. 'Why don't they just tell me straight, why allocate it?' The struggle to make sense of participating in a randomised controlled trial. Soc Sci Med 2002 Sep;55(5):709-719. [doi: 10.1016/s0277-9536(01)00197-6] [Medline: 12190265]
17. Denson A, Mahipal A. Participation of the elderly population in clinical trials: barriers and solutions. Cancer Control 2014 Jul;21(3):209-214. [doi: 10.1177/107327481402100305] [Medline: 24955704]
18. Lowe J, Archibald C. Cultural diversity: the intention of nursing. Nurs Forum 2009;44(1):11-18. [doi: 10.1111/j.1744-6198.2009.00122.x] [Medline: 19187049]
19. Mullins C, Abdulhalim A, Lavallee D. Continuous patient engagement in comparative effectiveness research. J Am Med Assoc 2012 Apr 18;307(15):1587-1588. [doi: 10.1001/jama.2012.442] [Medline: 22511684]
20. Graffigna G, Barello S. Patient Health Engagement (PHE) model in enhanced recovery after surgery (ERAS): monitoring patients' engagement and psychological resilience in minimally invasive thoracic surgery. J Thorac Dis 2018 Mar;10(Suppl 4):S517-S528 [FREE Full text] [doi: 10.21037/jtd.2017.12.84] [Medline: 29629198]

21. Burlew A, Weekes J, Montgomery L, Feaster D, Robbins M, Rosa C, et al. Conducting research with racial/ethnic minorities: methodological lessons from the NIDA Clinical Trials Network. Am J Drug Alcohol Abuse 2011 Sep;37(5):324-332 [FREE Full text] [doi: 10.3109/00952990.2011.596973] [Medline: 21854274]
22. Domecq J, Prutsky G, Elraiyah T, Wang Z, Nabhan M, Shippee N, et al. Patient engagement in research: a systematic review. BMC Health Serv Res 2014 Feb 26;14:89 [FREE Full text] [doi: 10.1186/1472-6963-14-89] [Medline: 24568690]
23. Wallerstein N, Duran B. Using community-based participatory research to address health disparities. Health Promot Pract 2006 Jul;7(3):312-323. [doi: 10.1177/1524839906289376] [Medline: 16760238]
24. Unertl K, Schaefbauer C, Campbell T, Senteio C, Siek K, Bakken S, et al. Integrating community-based participatory research and informatics approaches to improve the engagement and health of underserved populations. J Am Med Inform Assoc 2016 Jan;23(1):60-73 [FREE Full text] [doi: 10.1093/jamia/ocv094] [Medline: 26228766]
25. Library of Congress. All Information (Except Text) for HR34 - 21st Century Cures Act URL: https://www.congress.gov/bill/114th-congress/house-bill/34/all-info?r=3 [accessed 2018-12-14]
26. American Association for Cancer Research. National Cancer Moonshot Initiative URL: https://www.aacr.org/AboutUs/Pages/national-cancer-moonshot-initiative.aspx [accessed 2018-12-14]
27. Singer D, Jacks T, Jaffee E. A US 'Cancer Moonshot' to accelerate cancer research. Science 2016 Sep 9;353(6304):1105-1106. [doi: 10.1126/science.aai7862] [Medline: 27605537]
28. Rosenfield S, Kennedy A, Palmieri D, Bennett L. Abstract B51: Advancing cancer health disparities research through the Beau Biden Cancer MoonshotSM. In: Proceedings of the Tenth AACR Conference on The Science of Cancer Health Disparities in Racial/Ethnic Minorities and the Medically Underserved. 2017 Presented at: AACR'17; September 25-28, 2017; Atlanta, GA URL: https://cebp.aacrjournals.org/content/27/7_Supplement/B51
29. National Cancer Institute. Establish a Network for Direct Patient Engagement URL: https://www.cancer.gov/research/key-initiatives/moonshot-cancer-initiative/implementation/patient-engagement [accessed 2018-12-14]
30. Forsythe L, Ellis L, Edmundson L, Sabharwal R, Rein A, Konopka K, et al. Patient and stakeholder engagement in the PCORI pilot projects: description and lessons learned. J Gen Intern Med 2016 Jan;31(1):13-21 [FREE Full text] [doi: 10.1007/s11606-015-3450-z] [Medline: 26160480]
31. Frank L, Forsythe L, Ellis L, Schrandt S, Sheridan S, Gerson J, et al. Conceptual and practical foundations of patient engagement in research at the patient-centered outcomes research institute. Qual Life Res 2015 May;24(5):1033-1041 [FREE Full text] [doi: 10.1007/s11136-014-0893-3] [Medline: 25560774]
32. Henderson V, Barr K, An L, Guajardo C, Newhouse W, Mase R, et al. Community-based participatory research and user-centered design in a diabetes medication information and decision tool. Prog Community Health Partnersh 2013;7(2):171-184 [FREE Full text] [doi: 10.1353/cpr.2013.0024] [Medline: 23793248]
33. Hibbard J, Peters E. Supporting informed consumer health care decisions: data presentation approaches that facilitate the use of information in choice. Annu Rev Public Health 2003;24:413-433. [doi: 10.1146/annurev.publhealth.24.100901.141005] [Medline: 12428034]
34. Watson BR, Cavanah S. Community Information Needs: A Theory and Methodological Framework. Mass Commun Soc 2015;18(5):651-673 [FREE Full text] [doi: 10.1080/15205436.2015.1059948]
35. Shippee N, Garces JP, Lopez GJ, Wang Z, Elraiyah T, Nabhan M, et al. Patient and service user engagement in research: a systematic review and synthesized framework. Health Expect 2015 Oct;18(5):1151-1166 [FREE Full text] [doi: 10.1111/hex.12090] [Medline: 23731468]
36. National Cancer Institute. Methodology Reports URL: https://hints.cancer.gov/data/methodology-reports.aspx [accessed 2018-12-16]
37. National Cancer Institute. Health Information National Trends Survey URL: https://hints.cancer.gov/ [accessed 2019-08-12]
38. Levy BS, Sidel VW. Social Injustice and Public Health. Oxford, United Kingdom: Oxford University Press; 2013.
39. Dilworth-Anderson P, Pierre G, Hilliard T. Social justice, health disparities, and culture in the care of the elderly. J Law Med Ethics 2012;40(1):26-32. [doi: 10.1111/j.1748-720X.2012.00642.x] [Medline: 22458459]
40. Bellera C, Praud D, Petit-Monéger A, McKelvie-Sebileau P, Soubeyran P, Mathoulin-Pélissier S. Barriers to inclusion of older adults in randomised controlled clinical trials on Non-Hodgkin's lymphoma: a systematic review. Cancer Treat Rev 2013 Nov;39(7):812-817. [doi: 10.1016/j.ctrv.2013.01.007] [Medline: 23473865]
41. Dimond E, Germain D, Nacpil LM, Zaren HA, Swanson SM, Minnick C, et al. Creating a 'culture of research' in a community hospital: Strategies and tools from the National Cancer Institute Community Cancer Centers Program. Clin Trials 2015 Jun;12(3):246-256 [FREE Full text] [doi: 10.1177/1740774515571141] [Medline: 25691600]
42. Rahman S, Majumder M, Shaban S, Rahman N, Ahmed M, Abdulrahman K, et al. Physician participation in clinical research and trials: issues and approaches. Adv Med Educ Pract 2011;2:85-93 [FREE Full text] [doi: 10.2147/AMEP.S14103] [Medline: 23745079]
43. Vedsted P, Olesen F. Are the serious problems in cancer survival partly rooted in gatekeeper principles? An ecologic study. Br J Gen Pract 2011 Aug;61(589):e508-e512. [doi: 10.3399/bjgp11X588484] [Medline: 21801563]
44. Butow P, Kazemi J, Beeney L, Griffin A, Dunn S, Tattersall M. When the diagnosis is cancer: patient communication experiences and preferences. Cancer 1996 Jun 15;77(12):2630-2637. [doi: 10.1002/(SICI)1097-0142(19960615)77:12<2630::AID-CNCR29>3.0.CO;2-S] [Medline: 8640715]

45. George G, Iwuanyanwu E, Buford A, Piha-Paul S, Subbiah V, Fu S, et al. Cancer-related internet use and its association with patient decision making and trust in physicians among patients in an early drug development clinic: a questionnaire-based cross-sectional observational study. J Med Internet Res 2019 Mar 14;21(3):e10348 [FREE Full text] [doi: 10.2196/10348] [Medline: 30869638]
46. Rogers ZF, Gould SJ. How do you know that? The epistemology of consumer health decision making under conditions of risk-benefit conflict. Psychol Mark 2015 Mar 10;32(4):450-466 [FREE Full text] [doi: 10.1002/mar.20791]
47. Institute of Medicine, Committee on Quality of Health Care in America. Crossing the Quality Chasm: A New Health System for the 21st Century. Washington DC: National Academies Press; 2001.
48. Hesse B, Greenberg A, Peterson E, Chou WS. The health information national trends survey (HINTS): a resource for consumer engagement and health communication research. Stud Health Technol Inform 2017;240:330-346. [doi: 10.3233/ISU-170830] [Medline: 28972526]
49. Prey J, Woollen J, Wilcox L, Sackeim A, Hripcsak G, Bakken S, et al. Patient engagement in the inpatient setting: a systematic review. J Am Med Inform Assoc 2014;21(4):742-750 [FREE Full text] [doi: 10.1136/amiajnl-2013-002141] [Medline: 24272163]
50. LeRouge C, van Slyke C, Seale D, Wright K. Baby boomers' adoption of consumer health technologies: survey on readiness and barriers. J Med Internet Res 2014 Sep 8;16(9):e200 [FREE Full text] [doi: 10.2196/jmir.3049] [Medline: 25199475]
51. Tennant B, Stellefson M, Dodd V, Chaney B, Chaney D, Paige S, et al. eHealth literacy and Web 2.0 health information seeking behaviors among baby boomers and older adults. J Med Internet Res 2015 Mar 17;17(3):e70 [FREE Full text] [doi: 10.2196/jmir.3992] [Medline: 25783036]
52. Millery M, Ramos W, Lien C, Aguirre A, Kukafka R. Design of a community-engaged health informatics platform with an architecture of participation. AMIA Annu Symp Proc 2015;2015:905-914 [FREE Full text] [Medline: 26958227]
53. Weng C, Li Y, Ryan P, Zhang Y, Liu F, Gao J, et al. A distribution-based method for assessing the differences between clinical trial target populations and patient populations in electronic health records. Appl Clin Inform 2014;5(2):463-479 [FREE Full text] [doi: 10.4338/ACI-2013-12-RA-0105] [Medline: 25024761]
54. He Z, Carini S, Sim I, Weng C. Visual aggregate analysis of eligibility features of clinical trials. J Biomed Inform 2015 Apr;54:241-255 [FREE Full text] [doi: 10.1016/j.jbi.2015.01.005] [Medline: 25615940]

Abbreviations

BRP: Blue Ribbon Panel
CER: clinical effectiveness research
HINTS: Health Information National Trends Survey
NIH: National Institutes of Health
OR: odds ratio
PCORI: Patient-Centered Outcomes Research Institute

Edited by G Eysenbach; submitted 17.06.19; peer-reviewed by X Liu, G Kolostoumpis; comments to author 03.08.19; revised version received 08.09.19; accepted 24.09.19; published 29.10.19.

Please cite as:
Gerido LH, Tang X, Ernst B, Langford A, He Z
Patient Engagement in Medical Research Among Older Adults: Analysis of the Health Information National Trends Survey
J Med Internet Res 2019;21(10):e15035
URL: http://www.jmir.org/2019/10/e15035/
doi:10.2196/15035
PMID:31663860

Original Paper

Public Concern About Monitoring Twitter Users and Their Conversations to Recruit for Clinical Trials: Survey Study

Katja Reuter[1,2*], PhD; Yifan Zhu[3*], MLIS; Praveen Angyan[1], MS; NamQuyen Le[1], MPH; Akil A Merchant[4], MD; Michael Zimmer[5*], PhD

[1]Southern California Clinical and Translational Science Institute, Keck School of Medicine, University of Southern California, Los Angeles, CA, United States

[2]Institute for Health Promotion and Disease Prevention Research, Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA, United States

[3]School of Information Studies, University of Wisconsin-Milwaukee, Milwaukee, WI, United States

[4]Cedars Sinai Medical Center, Los Angeles, CA, United States

[5]Department of Computer Science, Marquette University, Milwaukee, WI, United States

[*]these authors contributed equally

Corresponding Author:
Katja Reuter, PhD
Southern California Clinical and Translational Science Institute
Keck School of Medicine
University of Southern California
2250 Alcazar Street # 200
Los Angeles, CA, 90089
United States
Phone: 1 3234422046
Fax: 1 3234422082
Email: katja.reuter@gmail.com

Abstract

Background: Social networks such as Twitter offer the clinical research community a novel opportunity for engaging potential study participants based on user activity data. However, the availability of public social media data has led to new ethical challenges about respecting user privacy and the appropriateness of monitoring social media for clinical trial recruitment. Researchers have voiced the need for involving users' perspectives in the development of ethical norms and regulations.

Objective: This study examined the attitudes and level of concern among Twitter users and nonusers about using Twitter for monitoring social media users and their conversations to recruit potential clinical trial participants.

Methods: We used two online methods for recruiting study participants: the open survey was (1) advertised on Twitter between May 23 and June 8, 2017, and (2) deployed on TurkPrime, a crowdsourcing data acquisition platform, between May 23 and June 8, 2017. Eligible participants were adults, 18 years of age or older, who lived in the United States. People with and without Twitter accounts were included in the study.

Results: While nearly half the respondents—on Twitter (94/603, 15.6%) and on TurkPrime (509/603, 84.4%)—indicated agreement that social media monitoring constitutes a form of eavesdropping that invades their privacy, over one-third disagreed and nearly 1 in 5 had no opinion. A chi-square test revealed a positive relationship between respondents' general privacy concern and their average concern about Internet research (P<.005). We found associations between respondents' Twitter literacy and their concerns about the ability for researchers to monitor their Twitter activity for clinical trial recruitment (P=.001) and whether they consider Twitter monitoring for clinical trial recruitment as eavesdropping (P<.001) and an invasion of privacy (P=.003). As Twitter literacy increased, so did people's concerns about researchers monitoring Twitter activity. Our data support the previously suggested use of the *nonexceptionalist methodology* for assessing social media in research, insofar as social media-based recruitment does not need to be considered exceptional and, for most, it is considered preferable to traditional in-person interventions at physical clinics. The expressed attitudes were highly contextual, depending on factors such as the type of disease or health topic (eg, HIV/AIDS vs obesity vs smoking), the entity or person monitoring users on Twitter, and the monitored information.

Conclusions: The data and findings from this study contribute to the critical dialogue with the public about the use of social media in clinical research. The findings suggest that most users do not think that monitoring Twitter for clinical trial recruitment constitutes inappropriate surveillance or a violation of privacy. However, researchers should remain mindful that some participants might find social media monitoring problematic when connected with certain conditions or health topics. Further research should isolate factors that influence the level of concern among social media users across platforms and populations and inform the development of more clear and consistent guidelines.

(J Med Internet Res 2019;21(10):e15455) doi:10.2196/15455

KEYWORDS

AIDS; cancer; clinical research; clinical trial; crowdsourcing; ethics; HIV; HPV; infoveillance; infodemiology; informed consent; Internet; research ethics; Mechanical Turk; MTurk; monitoring; obesity; privacy; public opinion; recruitment; smoking; social media; social network; surveillance; TurkPrime; Twitter

Introduction

Background

The success of clinical trials depends on the enrollment of study participants, also referred to as research participant recruitment. Recruitment involves attracting and selecting suitable study participants. It can be conducted through different communication channels (eg, newspapers, radio, television, posters, brochures, email, and social media). Without their involvement, medical and scientific progress that benefits patients would be impossible [1-5]. A recent systematic review found that 76.1% (131/172) of randomized clinical trials discontinued due to poor recruitment [6]. There is an urgent need for innovative solutions to address the issue of underenrollment in clinical trials [1]. We wanted to assess the feasibility of using Twitter user data for enhancing clinical trial recruitment. There is growing interest in using social media data for research, which is also referred to as *infoveillance* [7,8] or *digital epidemiology* [9]. This type of social media monitoring uses insights from social media users' activity and conversations to learn more about their attitudes and behaviors. *Active recruitment* occurs when research team members approach and interact with specific individuals to enroll them in research on the basis of pre-existing knowledge of characteristics that would make them suitable candidates for particular clinical trials [10]. We hypothesized that users' data and their conversations derived from the social network Twitter could serve as a useful tool to identify and recruit potential participants for specific clinical trials.

In the context of the Internet and social media, user privacy is commonly considered a process of boundary management where individuals regulate disclosures in their social relationships through adjustments to the transmission and sharing of personal information online. In her theory of communication privacy management, Petronio argues that individuals are regularly engaged in decisions about disclosing or concealing private information within any given context [11]. As the Internet and social media platforms become increasingly embedded into everyday life, they introduce new flows of information that challenge privacy norms and make managing boundaries more difficult. Such dynamism is central to the notion of networked privacy, which Marwick and Boyd define as the "ongoing negotiation of contexts in a networked ecosystem in which contexts regularly blur and collapse" [12]. Additionally, Nissenbaum's theory of contextual integrity takes context as its starting point [13]. The contextual integrity framework rests on the understanding that social interactions occur in particular contexts and that norms govern people's expectations of how personal information should flow within a given context. Rejecting the traditional dichotomy of public versus private information, as well as the notion that a user's preferences and decisions of privacy are independent of context, contextual integrity provides a framework for evaluating the flow of personal information between different agents; it also provides a framework for explaining why certain patterns of information flow might be acceptable in one context but viewed as problematic in another. These approaches to privacy on social media platforms prompted us to consider user expectations of appropriate information flows in the context of monitoring Twitter activity for the purpose of clinical trial recruitment.

Nearly one-quarter of American adults (22%) use Twitter [14]. Twitter users can send short messages, called *tweets*, that are limited to 280 characters [15]; they can also search for any public message and further engage with tweets (ie, they can *like*, *reply to*, and *retweet* [ie, share] them). Previous research suggested that Twitter provides a "rich and promising avenue for exploring how patients conceptualize and communicate about their specific health issues" [16] and provides an avenue for raising awareness of clinical trials and boosting enrollment [17-19].

To test the feasibility of Twitter monitoring for recruiting clinical trial participants, we decided to develop a use case for a multisite cancer study on acute myeloid leukemia (AML) with patients in remission. These patients present a uniquely challenging population to recruit for clinical studies. AML, when active, typically leads to severe symptoms and hospitalization. Hospitalized patients are more accessible to screen, identify, and recruit for clinical trials. Once AML patients have completed their consolidation chemotherapy, they only visit their doctor every 3-4 months. The clinical trial we chose for this case study was designed to recruit patients in the first 3 months after they complete their consolidation chemotherapy, precisely the time when these patients have only sporadic contact with the health care system. Traditional techniques employed during routine patient contact would not be possible for this population. Furthermore, since postremission maintenance therapy is not a routine part of clinical practice for AML, we were unlikely to receive referrals from community

physicians for this clinical trial. Therefore, we sought to examine the feasibility of a social media monitoring-enabled solution.

However, in their review of the study protocol, the Central Institutional Review Board (CIRB) of the National Institute of Cancer raised concerns about the potential breach of privacy using monitoring techniques on Twitter. The CIRB committee noted the following:

> *Those who openly share their information via social media platforms may still have an expectation of privacy and/or be unaware of the platform's privacy policies. To contact people after utilizing the approach of "active listening" may be perceived by some potential participants as eavesdropping on their conversations about their health... This may produce distrust and potential participants may interpret this as an invasion of their privacy even though social media is understood by many to be a public sphere. Privacy risks specific to [a Twitter user's] diagnosis may be increased by taking part in the study. The study team, by echoing the information about an individual's diagnosis, may amplify this information, so it's more likely to come to the attention of the public or an employer.*

We used this feedback as guidance and motivation for designing the following research study to ascertain people's attitudes and level of concern about the use of social media monitoring on Twitter for targeted clinical trial recruitment.

Study Objective and Hypotheses

Scientists have pointed out a lack of the inclusion of public views to inform future practices in social media research and social media-enabled recruitment [20,21]. Furthermore, a recent survey about the general use of tweets in research showed a lack of awareness among Twitter users that their public tweets could be used by researchers [22]. Therefore, the objective of this study was to examine the attitudes and level of concern among the public about using Twitter for the monitoring of social media users and their conversations in order to identify and recruit clinical trial participants. We focused on a variety of health topics, including cancer, obesity, human papilloma virus (HPV), HIV/AIDS, and smoking. The reason we chose to include a range of health topics, including nontransmissible and transmissible diseases, is that it reflects the spectrum of clinical trials that are being conducted in the United States and globally. We anticipated that the level of concern might vary by disease type and should be taken into consideration when choosing the recruitment method.

This study tested four primary and three additional hypotheses related to potential privacy concerns with the use of Twitter monitoring for clinical trial recruitment (see Textbox 1). Motivated by the CIRB's comments, we developed three hypotheses to test the CIRB's concerns regarding the potential for Twitter users to perceive social media monitoring as invasive and a violation of privacy: see Hypotheses 1-3 in Textbox 1. Drawing from Gelinas et al, we also sought to test the validity of the *nonexceptionalist methodology* [10], which suggests that recruiting clinical trial participants online should be normalized and should not be considered exceptional compared to traditional, offline recruitment strategies: see Hypothesis 4 in Textbox 1. They argue that "social media recruitment should be evaluated in substantially the same way as more traditional analogue or 'off-line' recruitment." Building from these four primary hypotheses, we sought to determine whether additional factors might impact participants' level of concern with social media monitoring for clinical trial participant recruitment. We isolated different factors within the vignettes for further analysis (ie, the type of information being monitored, the kind of disease or health topic of the clinical trial, and the nature of the entity engaged in the monitoring): see Hypotheses 5-7 in Textbox 1.

Our results are based on the views of the public and they support the formulation of evidence-based guidelines to assist researchers and Institutional Review Board (IRB) professionals using social media in clinical research recruitment. The data contribute to the critical dialogue with the public to understand the ethical issues involved in social media-enabled research and recruitment as well as the procedural solutions that are required to protect the rights and safety of research participants.

Textbox 1. Hypotheses we intended to test with this research study.

Hypothesis 1: People perceive social media monitoring on Twitter for clinical trial recruitment as eavesdropping on their conversations about their health and as an invasion of their privacy.
Hypothesis 2: Twitter users' expectations of privacy relate to their level of concern about the use of social media monitoring for clinical trial recruitment.
Hypothesis 3: General literacy and knowledge about the Twitter platform are associated with the level of concern about the use of social media monitoring on Twitter for clinical trial recruitment.
Hypothesis 4: People's concerns over Twitter monitoring for clinical trial recruitment are similar to more traditional, offline scenarios (eg, discretely being approached in person as the patient leaves a medical facility).
Hypothesis 5: The type of information monitored to identify and recruit individuals for clinical trials is associated with the level of concern over the use of social media monitoring on Twitter for clinical trial recruitment.
Hypothesis 6: The type of disease recruited for is associated with the level of concern over the use of social media monitoring on Twitter for clinical trial recruitment.
Hypothesis 7: The type of entity performing the monitoring is associated with the level of concern over the use of social media monitoring on Twitter for clinical trial recruitment.

Methods

Survey Instrument

We developed an open 39-item survey (see Multimedia Appendix 1) with the overall goal of assessing participants' attitudes and concerns regarding the use of Twitter for monitoring social media users and their conversations to identify and recruit clinical trial participants; we used a convenience sample. The following sections report on aspects of this survey study in accordance with the Checklist for Reporting Results of Internet E-Surveys (CHERRIES) [23]. Most questions were required; however, in some cases they were optional or allowed multiple answers. We incorporated two attention-check questions to assess respondents' attentiveness to the wording of questions and eliminated from the final dataset those respondents who failed them. We tested the survey in order to evaluate the reading level and complexity of the questions, acceptability of the instrument to participants, the respondent burden, and time needed to complete the instrument. Among the testers were two community members (promotora, ie, lay Hispanic or Latino community members who receive specialized training to provide basic health education in the community without being professional health care workers) from Los Angeles, three experts from the Community Engagement Core Group team at the Southern California Clinical and Translational Science Institute at the University of Southern California, and four graduate students from the University of Wisconsin-Milwaukee. We refined the survey instrument according to their feedback, in particular the wording of the vignettes and the true and false questions.

Using the survey, we collected the following types of information: previous use and knowledge of Twitter, general concern about Internet privacy, specific concerns about privacy related to the monitoring of Twitter activity for clinical trial recruitment, and demographic data. Clinical trials were defined for respondents in accordance with the National Institutes of Health definition for nonspecialist audiences [24]:

> *The goal of clinical trials is to determine if a new drug, device, or procedure works and is safe, or they can look at other aspects of care, such as improving the quality of life for people with chronic illnesses. People participate in clinical trials for a variety of reasons, for example, to help others [and] to contribute to moving science forward.*

Finally, we used a set of vignettes to assess the association between the level of concern and different variables, such as the disease or health topic of the clinical trial and the entity that monitors social media user activity on Twitter.

Participants

Eligible participants were adults, 18 years of age or older, who lived in the United States. People with and without Twitter accounts were included in the study.

Sample and Recruitment Methods

Overview

We used two online methods for recruiting study participants, who made up our convenience sample: the open survey was (1) advertised on Twitter between May 23 and June 8, 2017, and (2) deployed on TurkPrime, a crowdsourcing data acquisition platform, between May 23 and June 8, 2017 [25]. Accessing large numbers of participants from the Internet is referred to as crowdsourcing.

Twitter Recruitment

The Twitter ads appeared as promoted tweets in users' Twitter feeds. Twitter ads provide a number of targeting options for reaching a specific target audience. Targeting features used for ads in this study included (1) *age targeting* to adults aged 18 or older, (2) *location targeting* to the United States, (3) *language targeting* to users who understand English, and (4) *keyword and hashtag targeting* for words and hashtagged words that Twitter users have tweeted or searched for on Twitter related to four main categories. The four categories to which targeted keywords or hashtags were related were (1) social media and social media surveillance, (2) research participant recruitment and clinical trial enrolment, (3) ethics and Internet privacy, and (4) clinical research and clinical trials. Each ad included a brief description (eg, "Your opinion on social media surveillance on Twitter and for a chance at a gift card. Survey and raffle entry."), an image related to survey taking, a request for volunteers needed, a request for providing feedback, and a link to the questionnaire. Twitter ads were posted by the principal investigator's Twitter handle (ie, @dmsci). Our recruitment target was 500 participants. The daily maximum ad budget was set at US $49 with a total budget of US $980 for the entirety of the project. Respondents on Twitter had the opportunity to enter a raffle to win one of 10 US $100 gift cards upon completion of the survey. Duplicate and fraudulent responses were identified and removed as described by Teitcher et al [26]. More specifically, we used four methods to check for duplicate and fraudulent responses: (1) we checked for inconsistent and irregular answers, (2) we assessed the survey submission time stamps and batch submissions, (3) we examined email addresses that used random English words followed by three to six random letters (eg, upgradeyhujer@gmail.com), and (4) we contacted suspected respondents via email and asked them to verify the answers to three questions included in the survey to compare their responses (ie, their first name, age, and highest education level).

TurkPrime Recruitment

The second sample used in this study was recruited through TurkPrime [25], a panel service that allows researchers to target specific demographic groups. Prime Panels provides researchers with access to members of a number of market research panels through a Web interface similar to Amazon's crowdsourcing platform Mechanical Turk, which has been found to be an effective method to recruit study participants online across a wide spectrum of disciplines [27-34]. However, TurkPrime offers a proportional matching sampling approach. The study was visible to eligible participants on their dashboards. They

also received an email inviting them to participate in the study. We applied a *census-matched* template provided by TurkPrime that ensured that the sample proportionally matched the US adult population, aged 18 years or older, in terms of gender, age, race, ethnicity, and US region. More specifically, target benchmarks for key demographics included the following: gender—male (49.4%) and female (50.6%); age in years—18-29 (22.4%), 30-39 (16.8%), 40-49 (16.4%), 50-59 (17.8%), 60-69 (14.0%), and 70-99 (12.6%); Hispanic—not Hispanic (84.0%) and Hispanic, Latino, or Spanish (16.0%); and ethnicity—white (78.8%), black or African American (13%), American Indian or Alaska Native (1.2%), Asian (4.8%), and some other race (2.2%). These characteristics were targeted because they were underrepresented in the Twitter study convenience sample. Upon completion of the survey, study participants received compensation in the amount that they agreed to with the market research platform through which they entered the survey. Upon successful completion of the survey's attention-check questions, participants were also given bonuses. Bonuses serve as incentives for participation and have shown a substantial effect on data quality and the creativity of workers [35]. Target recruitment was 500 participants, for a total budget of US $3500. To ensure data protection, TurkPrime ensures the following [25]:

> *TurkPrime... uses transport layer security encryption (also known as HTTPS) for all transmitted data. All data access is blocked except for explicitly whitelisted IP addresses, in addition to being secured with user passwords. Furthermore, [the] data, including Access Key ID and Secret Access Key, are encrypted with AES-256 encryption, the standard adopted by the National Institute of Standards and Technology.*

Data Collection

Study data were collected and managed using Research Electronic Data Capture (REDCap), an electronic data capture tool, hosted at the University of Southern California. REDCap is a secure, Web-based application designed to support data capture for research studies, providing (1) an intuitive interface for validated data entry, (2) audit trails for tracking data manipulation and export procedures, (3) automated export procedures for seamless data downloads to common statistical packages, and (4) procedures for importing data from external sources [36].

The paid ads posted on Twitter included a link to the survey hosted on REDCap. Respondents filled out a multipage survey online on either a mobile device or desktop. On TurkPrime, each respondent was provided with a unique link to a separate survey hosted on REDCap. The datasets used for analysis were generated directly from REDCap using the platform's reporting tools. Please see the Twitter and TurkPrime recruitment sections for further details.

Data Cleaning

A total of 603 participants completed the survey and passed the attention-check questions in this study: 94 (15.6%) on Twitter and 509 (84.4%) on TurkPrime. Among the initial 704 respondents on Twitter alone, we used Excel filters to identify and remove 70 respondents (9.9%) who did not show correct completion of the attention-check questions and 540 respondents (76.7%) who gave fraudulent responses with unique characteristics. Regarding the fraudulent responses, all of them (1) showed the same age (ie, 22 years old); (2) were submitted about 5-10 minutes apart from each other over a period of 5 days; (3) used email addresses with a consistent pattern, namely, a random English word followed by three to six random letters (eg, upgradeyhujer@gmail.com and imageiunmed@gmail.com); and (4) were confirmed to be fraudulent when respondents were asked to verify the information provided through the survey about their first name, age, and highest degree or level of school they had completed. For each filtered entry, we manually reviewed the email address to identify fraudulent emails (ie, email addresses that included a random English word followed by three to six random letter patterns). Finally, we manually sent a message to each email address and asked the users to verify the information they provided in their survey responses. Among the initial 738 responses on TurkPrime, we removed 229 responses (31.0%) that did not show correct completion of the attention-check questions.

Data Analysis

We did not use any methods to adjust the sample, such as weighting of items or propensity scores. We analyzed the data on two levels: (1) at the respondent level to test control variables (individual factors: level 2) and (2) at the vignette level to test independent variables (contextual factors: level 1). Survey responses were first analyzed through descriptive statistical methods to assess the distribution of participants across our dependent and independent variables, such as the degree of privacy concern and demographic factors. Next, data regarding the different levels of concern for each vignette were further analyzed using pivot tables to identify any relationships between the levels of general privacy concern and the participants' attitudes regarding the vignettes. We defined a high level of concern as responses that indicated *Very or somewhat concerned* and a low level of concern as responses that indicated *Not too or not at all concerned*. Finally, we also analyzed the responses using descriptive and inferential statistical techniques (ie, crosstabs and chi-square tests) to determine, generally, where respondents had strong concerns regarding the use of Twitter monitoring in clinical trial recruitment and where respondents had a weaker understanding of Twitter's functions and usage policies. In particular, we looked to see if the concern regarding the use of Twitter monitoring in clinical trial recruitment was correlated with greater or less knowledge of Twitter, the type of monitoring used, or the method of outreach to the potential recruit. We report the results in aggregated form with all individually identifying information removed.

Institutional Review Board Review and Approval

The study was reviewed and approved by the IRB at the University of Southern California (HS-17-00348).

Results

Description of Participants

Demographics

Overall, the sample of 603 participants showed the following distribution (see Multimedia Appendix 2): 324 were male (53.7%) and 261 were female (43.3%); the majority were non-Hispanic white (421/603, 64.5%), 63 (9.7%) were Hispanic, and 66 (10.1%) were African American or black; and roughly a quarter (152/603, 25.2%) were 18-29 years of age and 107 (17.7%) were older than 60 years. The mean age of all respondents was 42.66 years (SD 16.00). Additionally, 151 (25.2%) respondents reported that their lives were affected by a chronic or rare disease.

Twitter Usage

We further assessed Twitter usage among the 603 survey participants (see Multimedia Appendix 3). Of the 603 respondents, 301 (49.9%) had a Twitter account at the time of the study, however, 300 valid responses were received for frequency and last time-usage questions, and 174 (28.9%) never used Twitter at all. A total of 186 out of 301 respondents (61.8%) who used Twitter had public accounts (ie, every Twitter user can view their account and messages), 199 out of 300 (66.3%) used the network at least weekly, 122 out of 300 (40.7%) used the network nearly every day, and more than half (181/300, 60.3%) had sent a Twitter message within the last week.

Twitter Literacy and Knowledge

We attempted to assess the level of Twitter literacy and knowledge among study participants (see Multimedia Appendix 4). Overall, 1209 of the total 3015 responses (40.10%) to the Twitter literacy questions that we collected from the 603 respondents were correct and 367 answers (12.17%) were incorrect, while nearly half of the responses (1439/3015, 47.73%) indicated that participants did not know. More specifically, 429 out of 603 respondents (71.1%) correctly answered when asked about the function of hashtags, while 138 (22.9%) did not know their function. When asked about Twitter account privacy settings, the majority of respondents (355/603, 58.9%) answered correctly, but 201 (33.3%) did not know about them. On the other hand, when asked about the automatic deletion of old Twitter messages after 1 year, 159 out of 603 respondents (26.4%) answered correctly and 385 (63.8%) did not know about this. When asked about the accessibility of public Twitter messages to unregistered Twitter visitors, 80 out of 603 respondents (13.3%) answered correctly, while 177 (29.4%) selected the wrong answer and 346 (57.4%) did not know the answer. Finally, when asked about Twitter's search capabilities that allow software programmers to search for Twitter messages by keyword and to collect profile information about the originating Twitter account, 186 out of 603 respondents (30.9%) answered correctly, while 369 (61.2%) did not know about these capabilities.

General Concern About Internet Privacy

We sought to learn more about general privacy concerns associated with the use of the Internet (see Multimedia Appendix 5). Of the 603 respondents, regardless of previous Twitter usage, 409 (67.8%) expressed some level of concern about their privacy while using the Internet. When asked how concerned respondents were about people they do not know obtaining personal information about them from their social media accounts and activities, 425 (70.5%) respondents expressed some level of concern. However, when asked how concerned respondents were about posts they made on social media that can be viewed by or shared with people not within their immediate network of friends or followers, fewer people (313/603, 51.9%) expressed some level of concern. As for these posts being used by companies for promotional purposes, 310 respondents (51.4%) expressed some level of concern. In contrast, 420 (69.7%) respondents expressed some level of concern about social media companies that might share or sell their information with third parties.

General Concern About Internet Research and Privacy

We also assessed respondents' concerns about Internet research activities, which pertain to the use of their Twitter data for research purposes (see Multimedia Appendix 6). We found that 252 of the 603 respondents (41.8%) expressed some level of concern regarding researchers' ability to send untargeted tweets visible to all their followers with a link for more information on how to participate in a clinical trial. Fewer respondents (226/603, 37.5%) expressed some level of concern about researchers noticing trending topics or hashtags related to health conditions, such as #Diabetes, #LungCancer, or #HeartDisease, and sending untargeted Twitter messages that include a link to more information on how to participate in a clinical trial, using the same hashtag. When asked how concerned they were about researchers actively monitoring users' Twitter activity to identify and contact potential participants for clinical trials based on the users' previous messages, 293 out of 603 respondents (48.6%) expressed some level of concern. However, fewer respondents (243/603, 40.3%) expressed some level of concern about researchers using paid Twitter advertisements (eg, *sponsored tweets*) to try to increase the likelihood that a clinical trial recruitment message gets seen by as many individuals as possible. Finally, 259 out of 603 respondents (43.0%) expressed some level of concern about Twitter keeping track of whether they clicked on a Twitter recruitment message related to a health study, for example, "Seeking participants for a #Cancer study."

Hypotheses Assessment

Hypothesis 1

Hypothesis 1 states that social media monitoring on Twitter for clinical trial recruitment is perceived as eavesdropping and an invasion of privacy.

To gauge respondents' overall perception of Twitter monitoring for clinical trial recruitment, we tested the language as stated by the CIRB that *active listening* may be perceived by participants as *eavesdropping* on their conversations about their health (see Multimedia Appendix 7). When asked about monitoring of public Twitter conversations by medical

researchers to identify and recruit potential clinical trial participants, 269 of 603 respondents (44.7%) considered it eavesdropping, while 333 (55.3%) did not consider it eavesdropping or did not know. Out of 603 respondents, 259 (43.0%) thought the monitoring was an invasion of their privacy, while 344 (57.0%) did not consider it an invasion of privacy or did not know. Finally, 235 of 603 respondents (39.0%) thought the monitoring was a potential breach of confidentiality, while 368 (61.1%) did not consider it a breach of confidentiality or did not know.

We isolated responses for only those respondents (409/603, 67.8%) who expressed some level of general concern about their privacy while using the Internet; we combined *Very concerned* with *Somewhat concerned* responses. These respondents' overall opinions regarding the questions about eavesdropping, privacy, and confidentiality revealed slightly greater privacy concerns than the entire population. As reported in Multimedia Appendix 8, out of 409 respondents, 199 (48.8%) considered Twitter monitoring as eavesdropping, 202 (49.4%) considered it an invasion of their privacy, and 180 (44.0%) thought that it could jeopardize confidentiality.

We also examined the responses of those participants (178/603, 29.5%) who expressed little or no general concern about their overall privacy while using the Internet; this allowed us to assess whether those with little general privacy concern might still have elevated privacy concern about Twitter monitoring. Those with lower general Internet privacy concern indicated lower concern in response to the questions about eavesdropping, privacy, and confidentiality (see Multimedia Appendix 8). Similarly, fewer respondents with active Twitter accounts (199/603, 33.0%) indicated concerns with Twitter monitoring compared to the overall population (see Multimedia Appendix 7).

Hypothesis 2

Hypothesis 2 states that the expectation of Internet privacy relates to the level of concern about Internet research and Twitter monitoring for clinical trial recruitment.

We wanted to gauge whether the presence of general Internet privacy concern is related to increased concern about Internet research (see Multimedia Appendix 9). Therefore, we isolated responses for only those respondents (409/603, 67.8%) who expressed some level of general concern about their privacy while using the Internet—we combined *Very concerned* with *Somewhat concerned* responses—and compared them to the entire population reported in Multimedia Appendix 6. These respondents showed higher levels of general Internet research privacy concern. For example, 235 out of 409 respondents (57.5%) indicated concern about researchers actively monitoring Twitter to identify and contact potential participants for clinical trials, compared to only 293 respondents out of the entire population of 603 (48.6%).

Isolating for only those respondents (178/603, 29.5%) who expressed little or no general privacy concern, we found that this population generally had lower levels of Internet research privacy concern (see Multimedia Appendix 9). For example, only 55 of 178 respondents (30.9%) indicated concern about researchers actively monitoring Twitter activity to identify and contact potential clinical trial participants, compared to 293 respondents out of the entire population of 603 (48.6%). Similarly, only 32 of 178 respondents (18.0%) showed concern about researchers' monitoring of hashtags in tweets, generally, compared to 244 respondents out of the entire population of 603 (40.5%) and 206 of the 409 respondents (50.4%) with high privacy concerns. A chi-square test was used to explore whether there is a relationship between respondents' general privacy concerns and their average concerns about Internet research. The test, taking into account the population of 603 participants, revealed a statistically significant relationship between these variables: $\chi^2_{16}=143.0$, $P<.005$. We then stratified responses based on Twitter use to assess whether active users of the social media platform expressed different levels of privacy concern regarding the use of Twitter for research purposes (see Multimedia Appendix 9). Respondents with active Twitter accounts (199/603, 33.0%) who indicated that they used the platform once a week or more reported lower levels of general Internet research privacy concern compared to the entire population. Our data suggest that being an active Twitter user might impact the levels of privacy concern expressed regarding Twitter-based Internet research activities.

Finally, we stratified the responses to the Twitter-monitoring vignettes (see Multimedia Appendix 10) based on respondents' overall levels of Internet privacy concern and whether they are active Twitter users. We analyzed each vignette's subquestions, isolating responses for those who expressed some concern and those who did not. Upon analyzing responses from the 409 participants out of 603 (67.8%) who expressed some level of general concern about their privacy while using the Internet, we discovered that a larger proportion of these respondents indicated some concern regarding each of the various Twitter-monitoring vignettes compared to the entire population (see Table 1).

We also isolated responses for those respondents (178/603, 29.5%) who expressed little or no general concern about their overall privacy while using the Internet; this allowed us to assess whether those with little general privacy concern might still have elevated privacy concern about the types of Twitter monitoring described in the vignettes. As reported in Table 1, those with lower general privacy concern indicated much lower concern over the vignette scenarios. Similarly, fewer respondents with active Twitter accounts (199/603, 33.0%) indicated concern with the Twitter-monitoring vignettes compared to the overall population, with a majority expressing concern only for the HIV/AIDS vignette. Overall, all groups expressed the most concern for the HIV/AIDS vignette and they expressed the least concern for the smoking vignette.

Finally, we performed chi-square tests to explore whether there was a relationship between general Internet privacy concern and levels of concern expressed with each vignette. The tests revealed a statistically significant relationship in all cases ($P<.001$), as reported in Table 2.

Table 1. Stratified analysis of vignette scenarios for respondents who indicated that they were Very concerned or Somewhat concerned about Twitter monitoring.

Vignette	Respondents (N=603), n (%)	Respondents with high general privacy concern (n=409), n (%)	Respondents with low general privacy concern (n=178), n (%)	Respondents who were active Twitter users (n=199), n (%)
Cancer vignette	300 (49.8)	244 (59.7)	51 (28.7)	75 (37.7)
Obesity vignette	299 (49.6)	241 (58.9)	52 (29.2)	76 (38.2)
HPV[a] vignette	298 (49.4)	243 (59.4)	51 (28.7)	75 (37.7)
HIV/AIDS vignette	349 (57.9)	269 (65.8)	73 (41.0)	106 (53.3)
Smoking vignette	255 (42.3)	207 (50.6)	45 (25.3)	66 (33.2)

[a]HPV: human papilloma virus.

Table 2. Chi-square analysis of concern expressed by respondents for each vignette based on their general privacy concern.

Vignette	Number of valid cases, N	Pearson chi-square	*df*	*P* (asymptotic significance, 2-sided)
Cancer vignette	603	175.9	16	<.001
Obesity vignette	603	126.7	16	<.001
HPV[a] vignette	603	124.4	16	<.001
HIV/AIDS vignette	603	79.6	16	<.001
Smoking vignette	603	102.5	16	<.001

[a]HPV: human papilloma virus.

Hypothesis 3

Hypothesis 3 states that general Twitter literacy is associated with the level of concern about the use of social media monitoring on Twitter for clinical trial recruitment.

There was a significant association (*P*=.001) between respondents' Twitter literacy and their concerns about the ability of researchers to monitor their Twitter activity, generally, for the purpose of clinical trial recruitment (see Table 3). This relationship also indicates that as Twitter literacy increases, so do people's concerns about researchers monitoring Twitter activity. Additionally, there was a significant association (*P*=.004) between respondents' Twitter literacy and their concerns about researchers monitoring particular information types on Twitter (eg, hashtags, public tweets, and profile description) for the purpose of clinical trial recruitment. Overall, there was a significant association (*P*=.03) between respondents' Twitter literacy and their overall concerns with researchers monitoring Twitter activity.

Related to the CIRB's concerns, we also found a significant association between respondents' Twitter literacy and whether they considered Twitter monitoring for clinical trial recruitment as eavesdropping (*P*<.001) and an invasion of privacy (*P*=.003). There was no significant association, however, between Twitter literacy and whether respondents felt that Twitter monitoring jeopardized confidentiality (*P*=.43).

Table 3. Chi-square analysis of concerns expressed by respondents based on their Twitter literacy.

Respondents' concerns	Number of valid cases, N	Pearson chi-square	*df*	*P* (asymptotic significance, 2-sided)
Concern about the ability for researchers to monitor their Twitter activity, generally	536	22.7	6	.001
Concern about researchers monitoring particular information types on Twitter (eg, hashtags, public tweets, and profile description)	556	19.3	6	.004
Overall concern with researchers monitoring Twitter activity	513	7.2	2	.03
Consider Twitter monitoring for clinical trial recruitment as eavesdropping	602	38.1	4	<.001
Consider Twitter monitoring for clinical trial recruitment as an invasion of privacy	603	15.8	4	.003
Felt Twitter monitoring jeopardizes confidentiality	603	3.9	4	.43

Hypothesis 4

Hypothesis 4 states that there are differences in attitudes toward Twitter monitoring for clinical trial recruitment compared to a more traditional, offline scenario.

We also used the vignettes to assess the attitudes toward a more traditional, offline scenario (see Multimedia Appendix 10). We asked participants about their attitudes toward patients discretely being approached in person as they leave a medical facility. We found that out of all 603 respondents, regardless of previous Twitter usage and across all disease types, fewer than one-third would be more comfortable with a traditional, in-person request to join a clinical trial: cancer (176/603, 29.2%), obesity (161/603, 26.7%), HPV (169/603, 28.0%), HIV/AIDS (174/603, 28.9%), and smoking (161/603, 26.7%). For the respondents with greater overall general Internet privacy concern, there was no meaningful shift in the respondents' comfort levels with having researchers recruit them as a research participant in person versus through Twitter monitoring.

Hypothesis 5

Hypothesis 5 states that the level of concern is associated with the type of information monitored for the purpose of identifying individuals to recruit for clinical trials.

We assessed the level of concern about the type of information medical researchers or research institutions might monitor and review in order to identify individuals for recruiting them into clinical trials (see Multimedia Appendix 6). When asked about monitoring of hashtags in tweets (ie, keywords used to organize and link conversations on Twitter, such as #SleepApnea, #Depression, or #HeartDisease), 244 of 603 respondents (40.5%) expressed some level of concern. When asked about reviewing the text of users' public Twitter messages, out of 603 respondents, 265 (43.9%) expressed some level of concern, while 285 (47.3%) expressed some level of concern about reviewing the text of their profile description.

Hypotheses 6 and 7

Hypotheses 6 and 7 state that there is a level of concern associated with the type of disease recruited for and the type of entity performing the monitoring.

We used the set of vignettes (see Table 1) to further assess the association between the level of concern and the disease or health topic of the clinical trial and the entity that monitors social media user activity on Twitter. We found that of all 603 respondents, regardless of previous Twitter usage, most people expressed some level of concern in response to the scenario of researchers at a medical research university monitoring for an HIV/AIDS trial (349/603, 57.9%). We compared this to respondents with some level of concern in response to other disease topics and entities, such as cancer and a research team at a major research institution (300/603, 49.7%), obesity and scientists at a pharmaceutical company (299/603, 49.6%), HPV vaccination and a health officer at a state public health office (298/603, 49.4%), and smoking and a health officer at a local public health office (255/603, 42.3%). For most vignettes, the type of entity that conducted the research was selected as the most important factor contributing to the level of concern; for example, for the cancer vignette, 284 out of 603 respondents (47.1%) indicated that the entity was the most important factor, while for the obesity vignette it was 286 respondents (47.4%), for the HPV vignette it was 271 respondents (44.9%), and for the HIV/AIDS vignette it was 250 respondents (41.5%).

We further stratified responses for each vignette's subquestions, isolating responses for those who expressed some concern—indicated *Very concerned* or *Somewhat concerned*—and those who expressed little or no concern—indicated *Not too concerned* or *Not concerned at all*—with the overall vignette scenario. As shown in Multimedia Appendix 11, *Who (or the entity who) is doing the Twitter monitoring* was the most common factor that impacted concern across all scenarios, regardless of whether the overall Internet privacy concern was low or high; the exception was with the HIV/AIDS scenario, where respondents who expressed overall concern noted that *The nature of the disease/medical condition being monitoredfor* was the main contributing factor. For the obesity and HPV scenarios, a noticeably larger portion of the respondents who expressed some concern also noted that the *Use of Twitter as a method in which the researchers contacted you* was also a contributing factor.

Data Availability

All relevant data that support the findings of this study are available in the data repository figshare:

1. Responses from Twitter users: Monitoring Twitter for clinical trial recruitment [37].
2. Responses from TurkPrime workers: Monitoring Twitter for clinical trial recruitment [38].

Discussion

Principal Findings

Public social networks such as Twitter provide access to user information, including personal and sensitive data, without necessarily requiring an individual's knowledge or consent. While previous studies explored the unique ethical challenges of social media as a health research tool and research data source [10,20,39,40], there are only a few studies that offer users' perspectives and public views on the use of social media monitoring as a clinical research recruitment tool [20,22]. For example, in a recent study, Fiesler et al found that the majority of surveyed Twitter users "felt that researchers should not be able to use tweets without consent" [22]. However, researchers have pointed out the need for views of the public on the subject to inform the development of ethical and regulatory guidelines and future practice [20,22].

The goal of this study was to contribute data that reflect public views of Twitter users and nonusers and to inform the scientific discourse about the use of Twitter user data for clinical trial recruitment. We discuss our findings in relation to our hypotheses (see Table 4) and contextual factors (eg, monitored information, study disease type, and monitoring entity) and conclude with potential implications for the practice.

Table 4. Summary of study findings by study hypothesis.

Hypotheses	Overall findings (nonstratified)
Primary hypotheses: derived from CIRB[a] feedback	
Hypothesis 1: Social media monitoring on Twitter for clinical trial recruitment is perceived as eavesdropping and as an invasion of privacy.	Not supported. While nearly half the respondents indicated agreement that social media monitoring constitutes a form of eavesdropping that invades their privacy, over one-third disagreed and nearly 1 in 5 had no opinion. Fewer respondents felt that social media monitoring jeopardizes confidentiality.
Hypothesis 2: Twitter users' expectations of privacy relate to their level of concern about the use of social media monitoring for clinical trial recruitment.	Supported. Chi-square tests revealed a positive relationship between respondents' general privacy concerns and their average concerns about Internet research (N=603): χ^2_{16}=143.0, *P*<.005. Additionally, respondents who indicated some general privacy concern also generally expressed greater concern over social media monitoring, in general, as well as for each vignette scenario. Chi-square tests confirmed a statistically significant relationship between general privacy concern and concern for each vignette.
Hypothesis 3: General literacy about the Twitter platform is associated with the level of concern about the use of social media monitoring on Twitter for clinical trial recruitment.	Supported. There was a statistically significant association (*P*=.001) between respondents' Twitter literacy and their concerns about the ability for researchers to monitor their Twitter activity, generally, for the purpose of clinical trial recruitment. Overall, as Twitter literacy increased, so did people's concerns about researchers monitoring Twitter activity. While there was an association between respondents' Twitter literacy and whether they consider Twitter monitoring for clinical trial recruitment as eavesdropping or an invasion of privacy, there was no significant association with whether respondents felt Twitter monitoring jeopardizes confidentiality.
Testing the validity of the *nonexceptionalist methodology*	
Hypothesis 4: People's concerns over Twitter monitoring for clinical trial recruitment are similar to those of more traditional, offline scenarios (eg, discretely approaching a patient in person as they leave a medical facility).	Supported. Most people were either indifferent, did not know, or were less comfortable with an in-person approach, regardless of previous Twitter usage and across all disease types. They did not find Twitter monitoring any more concerning than the more traditional means of clinical trial subject recruitment. Overall, the data presented here support the use of the *nonexceptionalist methodology* for assessing social media-based monitoring and recruitment.
Factors that might impact the level of concern over social media monitoring for clinical trial recruitment	
Hypothesis 5: The type of information monitored for the purpose of identifying individuals to recruit for clinical trials is associated with the level of concern over the use of social media monitoring on Twitter for clinical trial recruitment.	Partially supported. While not a majority, nearly half the respondents did indicate general concern about researchers actively monitoring users' Twitter activity to identify and contact potential participants for clinical trials. The greatest concern was related to reviewing the text of their profile description, with less concern expressed related to monitoring hashtags or the text of individual tweets.
Hypothesis 6: The type of disease recruited for is associated with the level of concern over the use of social media monitoring on Twitter for clinical trial recruitment.	Supported. Nearly 6 out of 10 respondents expressed concern about monitoring for an HIV/AIDS trial compared to other disease topics that raised less concern, such as cancer, obesity, HPV[b] vaccination, and smoking.
Hypothesis 7: The nature of the entity performing social media monitoring on Twitter is associated with the level of concern over this monitoring for clinical trial recruitment.	Supported. The factor that most impacted the level of concern was the entity or person who conducted the Twitter monitoring and research. The exception was the HIV/AIDS scenario, where respondents who expressed overall concern noted that *The nature of the disease/medical condition being monitored for* was the main contributing factor.

[a]CIRB: Central Institutional Review Board.

[b]HPV: human papilloma virus.

The Central Institutional Review Board's Concerns

When we tested the concerns raised by the CIRB that *active listening* may be perceived by participants as *eavesdropping* on their conversations about their health, an invasion of their privacy, and a potential breach of confidentiality, we found that the majority of respondents did not share this view. While the CIRB's concerns have some basis, with 4 in 10 respondents feeling Twitter monitoring is eavesdropping and an invasion of privacy, the concern was not widespread, even among those expressing higher levels of general online privacy concern. This suggests that while clinical researchers should be mindful that some Twitter users will be wary of being monitored for the purpose of clinical trial recruitment, these concerns should not

prevent the recruitment strategy from being pursued. Tactics such as Privacy by Design [21], for example, through privacy notices and disclaimers, can be applied to achieve privacy in social media-based research recruitment. Our data also show a statistically significant relationship between respondents' general privacy concern and their average concern about Internet-based research activities. Those who were generally more concerned about Internet privacy were also more concerned about different aspects of Twitter monitoring for trial recruitment, such as who was performing the monitoring and what information was being monitored. We found the opposite effect among those respondents who were generally less concerned about Internet privacy and who were active, frequent Twitter users. Our data suggest that being an active Twitter user might impact the level of privacy concern expressed regarding Twitter-based Internet research activities. This suggests that users who are more active online and aware of general privacy concerns are also more likely to be concerned about Twitter monitoring for clinical trial recruitment, due to a higher overall awareness of privacy and surveillance online.

Furthermore, the CIRB committee noted that "those who openly share their information via social media platforms may still be unaware of the platforms' privacy policies." We found that there is a significant association between respondents' Twitter literacy and their concerns about the ability for researchers to monitor their Twitter activity, generally, for the purpose of clinical trial recruitment. We further found a significant association between respondents' Twitter literacy and their concerns about researchers monitoring particular information types on Twitter (eg, hashtags, public tweets, and profile description) for the purpose of clinical trial recruitment. We cannot state, however, that these concerns necessarily increase as Twitter literacy increases. Related to the CIRB's concerns, we also found a significant association between respondents' Twitter literacy and whether they consider Twitter monitoring for clinical trial recruitment as eavesdropping and an invasion of privacy; however, there was no significant association between Twitter literacy and whether respondents felt Twitter monitoring jeopardizes confidentiality. Overall, there is a significant association between respondents' Twitter literacy and their overall concern with researchers monitoring Twitter activity, suggesting that the more that users understood about Twitter as a platform, the greater they were concerned about researchers monitoring their Twitter activity. This presents a challenge seen in many areas of online literacy, as confirmed in studies of Internet users, in general [41,42], as well as with social network users, in particular [22,43]. Thus, on the one hand, the more that people understand social media platforms, the more they are aware of possible privacy concerns. On the other hand, those who do not have high Twitter literacy might not be expressing concerns because they simply do not understand the potential threat.

Testing the Nonexceptionalist Methodology

Gelinas et al suggested employing a *nonexceptionalist methodology* for assessing social media recruitment in research and "normalizing social media recruitment techniques while remaining sensitive to their potentially novel aspects" [10]. They argue that "social media recruitment should be evaluated in substantially the same way as more traditional analogue or 'off-line' recruitment." This includes (1) the identification of "a more familiar off-line variant or equivalent of the social media technique being proposed," (2) identification of substantive ethical considerations with a focus on the respect for the privacy and other interests of social media users and investigator transparency, and (3) clarification and evaluation of any aspects in which the online version differs from the more traditional offline equivalent. We used a series of vignettes to assess respondents' attitudes toward a more traditional, offline scenario and asked them about their attitudes toward patients discretely being approached in person as they leave a medical facility. We found that, regardless of previous Twitter usage and across all disease types, most people were either indifferent, did not know, or were less comfortable with an in-person approach. This suggests that even while many respondents expressed concern over social media monitoring as eavesdropping or a potential violation of privacy, as noted above, they did not find it any more concerning than the more traditional means of clinical trial subject recruitment. In fact, our data show that less than one-third of the respondents preferred in-person recruitment over the Twitter-monitoring approach described in the vignettes. Even among those with a high level of general online privacy concern, only 38% preferred in-person recruitment. However, in-person recruitment is the current standard practice. Our findings support Gelinas et al, insofar as social media-based recruitment in itself does not need to be considered exceptional from the participant's perspective, while researchers should also remain mindful that some participants will find it problematic.

Additional Factors That Influence the Level of Concern

Following Marwick and Boyd [12] and Nissenbaum [13], our findings support the notion that users frame privacy concerns in online platforms contextually and that when contexts collapse or blur, privacy concerns might emerge. Nearly half the respondents indicated general concern about researchers actively monitoring users' Twitter activity to identify and contact potential participants for clinical trials. This suggests that, for many, a context collapse occurred that triggered some level of privacy concern; for example, information posted publicly for one reason, such as to share with one's Twitter followers, was taken from that social context and used for a different purpose (ie, clinical trial recruitment).

Our findings further support the point previously made by Bender et al [21] that "within health information, there are gradients of sensitivity," and certain health topics and disease types, such as cancer, may be considered less-sensitive personal health information. We found that the monitoring of Twitter user data that was related to HIV/AIDS raised the highest level of concern compared to monitoring related to cancer, HPV, obesity, or smoking. This may be partly due to the fact that HIV/AIDS is still associated with stigma [44]. Survey respondents commented as follows:

> *HIV a very serious and private disease... it is something that needs to be discussed in person.*

> *On Twitter, users are using the specific language. These users have already disclosed their opinions or*

> *diagnosis. I feel like it's similar to outing someone on accident if a company were to just randomly ask people.*

However, respondents also argued in favor of using Twitter monitoring for clinical trial recruitment:

> *If you talk about HIV/AIDS on Twitter or any social media, you have to know it's not private.*

> *As long as the person or researcher making contact with the target is being very transparent about the source of the research and is happy to give information to verify their identity and intent, I wouldn't be alarmed or put off.*

We identified additional factors that influenced the level of concern about monitoring Twitter user data for clinical trial recruitment. With the exception of the HIV/AIDS scenario as stated above, the factor that most impacted the level of concern was the type of entity or the person who conducted the research. Researchers who may use this approach should ensure investigator transparency; for example, investigators should refrain from fabricating online identities and clearly disclose the goal and design of the research [10]. In the case of monitoring Twitter user data for clinical trial recruitment, multiple messages could be used to introduce the project and main purpose of the outreach, as described by Reuter et al [45].

Finally, the form of contact on Twitter (ie, public replies versus private messages) played a more important role for the HIV/AIDS, obesity, and HPV scenarios, where a noticeably larger portion of the respondents expressed some concern. Respondents argued as follows:

> *This condition definitely need[s] to be addressed privately and not through a public reply.* [Participant in response to the HIV/AIDS vignette]

> *I think the public reply instead of a dm [direct message] could be embarrassing.* [Participant in response to the obesity vignette]

> *The nature of this can be very embarrassing and a public reply could be damaging.* [Participant in response to the HPV vignette]

This may be due to the stigma [44] associated with a disease such as HIV/AIDS and obesity or the level of controversy around a topic such as vaccination [46]. See Multimedia Appendix 12 for a broader sample of respondents' comments in response to vignettes.

Study Limitations

This study was limited to two populations: Twitter users and TurkPrime workers. The range of ages, education levels, and socioeconomic statuses of these populations could be more limited than those found in the general public. A total of 22% of US adults use Twitter, nearly equally among white, black, and Hispanic adults across all ages but with the highest usage among those 18-29 years of age [14]. TurkPrime workers (ie, turkers) are diverse across several demographic dimensions, such as age, gender, and income, but are not precisely representative of the United States as a whole [47]. Therefore, our findings may also not be generalizable to the monitoring of other social media platforms with different norms and privacy expectations, such as Reddit, Facebook, Instagram, Tumblr, or Snapchat. Although we expect to see similarities in public attitudes, future research will need to shed more light on how the results presented here might play out across different populations and different platforms.

Additionally, this was an exploratory study prompted by the feedback from a national research organization (ie, CIRB) and the sample size of this study was limited. More robust studies with a larger sample could yield additional insights. Finally, we acknowledge that while we chose seven hypotheses for this initial study, there are certainly other issues and variables that deserve further attention related to the subject in future studies.

Conclusions

The data we presented here contribute to the critical dialogue with the public about the use of social media in clinical research. Public social networks such as Twitter offer the clinical research community a novel opportunity for identifying and engaging potential study participants based on user activity data. However, the availability of public social media data has led to new ethical challenges about respecting user privacy and the appropriateness of monitoring social media for clinical trial recruitment. The results of this study suggest that most users do not think monitoring Twitter for the purpose of clinical trial recruitment constitutes inappropriate surveillance or a violation of privacy. Our data further support the previously suggested use of the *nonexceptionalist methodology* for assessing social media in research, insofar as social media-based recruitment in itself does not need to be considered exceptional from the participant's perspective and, for most, it is considered preferable to traditional in-person interventions at physical clinics. Notwithstanding these findings, researchers should also remain mindful that some participants might find social media monitoring problematic when connected with certain conditions. The expressed attitudes were highly contextual, depending on factors such as the type of disease or health topic and the entity or person who monitored users on Twitter. Further research should isolate factors that influence the level of concern among social media users across platforms and inform the development of more clear and consistent guidelines.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Open 39-item survey used in this study.

[PDF File (Adobe PDF File), 185 KB - jmir_v21i10e15455_app1.pdf]

Multimedia Appendix 2
Respondents' demographics.
[PDF File (Adobe PDF File), 101 KB - jmir_v21i10e15455_app2.pdf]

Multimedia Appendix 3
Respondents' Twitter usage.
[PDF File (Adobe PDF File), 49 KB - jmir_v21i10e15455_app3.pdf]

Multimedia Appendix 4
Respondents' Twitter literacy.
[PDF File (Adobe PDF File), 51 KB - jmir_v21i10e15455_app4.pdf]

Multimedia Appendix 5
Respondents' general Internet privacy concerns.
[PDF File (Adobe PDF File), 54 KB - jmir_v21i10e15455_app5.pdf]

Multimedia Appendix 6
Respondents' Internet research privacy concerns.
[PDF File (Adobe PDF File), 64 KB - jmir_v21i10e15455_app6.pdf]

Multimedia Appendix 7
Respondents' overall opinion of social media listening on Twitter for clinical trial recruitment.
[PDF File (Adobe PDF File), 81 KB - jmir_v21i10e15455_app7.pdf]

Multimedia Appendix 8
Stratified analysis of the overall opinion of social media listening on Twitter for clinical trial recruitment for those responses that indicated agreement with the concerning issues of eavesdropping, invasion of privacy, and jeopardized confidentiality.
[PDF File (Adobe PDF File), 12 KB - jmir_v21i10e15455_app8.pdf]

Multimedia Appendix 9
Stratified analysis of respondents who indicated "Very concerned" or "Somewhat concerned" about their privacy while using the Internet.
[PDF File (Adobe PDF File), 54 KB - jmir_v21i10e15455_app9.pdf]

Multimedia Appendix 10
Respondents' levels of concern about scenarios described in vignettes with a focus on study disease and the entity monitoring the social media user activity.
[PDF File (Adobe PDF File), 141 KB - jmir_v21i10e15455_app10.pdf]

Multimedia Appendix 11
Responses to vignette subquestions, stratified based on level of overall Internet privacy concern with vignette scenario.
[PDF File (Adobe PDF File), 96 KB - jmir_v21i10e15455_app11.pdf]

Multimedia Appendix 12
A select sample of respondents' comments in response to the vignettes.
[PDF File (Adobe PDF File), 128 KB - jmir_v21i10e15455_app12.pdf]

References

1. Bower P, Brueton V, Gamble C, Treweek S, Smith CT, Young B, et al. Interventions to improve recruitment and retention in clinical trials: A survey and workshop to assess current practice and future priorities. Trials 2014 Oct 16;15:399 [FREE Full text] [doi: 10.1186/1745-6215-15-399] [Medline: 25322807]
2. Treweek S, Pitkethly M, Cook J, Fraser C, Mitchell E, Sullivan F, et al. Strategies to improve recruitment to randomised trials. Cochrane Database Syst Rev 2018 Feb 22;2:MR000013. [doi: 10.1002/14651858.MR000013.pub6] [Medline: 29468635]

3. Institute of Medicine. Workshop summary. In: Proceedings of the Forum on Drug Discovery, Development, and Translation: Envisioning a Transformed Clinical Trials Enterprise in the United States: Establishing an Agenda for 2020. Washington, DC: National Academies Press; 2012 Presented at: Forum on Drug Discovery, Development, and Translation: Envisioning a Transformed Clinical Trials Enterprise in the United States: Establishing an Agenda for 2020; November 7-8, 2011; Washington, DC URL: https://www.nap.edu/catalog/13345/envisioning-a-transformed-clinical-trials-enterprise-in-the-united-states [doi: 10.17226/13345]
4. Institute of Medicine. Workshop summary. In: Proceedings of the Forum on Drug Discovery, Development, and Translation: Public Engagement and Clinical Trials: New Models and Disruptive Technologies. Washington, DC: National Academies Press; 2012 Presented at: Forum on Drug Discovery, Development, and Translation: Public Engagement and Clinical Trials: New Models and Disruptive Technologies; June 27-28, 2011; New York, NY URL: https://www.ncbi.nlm.nih.gov/books/NBK91498/ [doi: 10.17226/13237]
5. Fonseca DA, Amaral I, Pinto AC, Cotrim MD. Orphan drugs: Major development challenges at the clinical stage. Drug Discov Today 2019 Mar;24(3):867-872. [doi: 10.1016/j.drudis.2019.01.005] [Medline: 30658132]
6. Briel M, Olu KK, von Elm E, Kasenda B, Alturki R, Agarwal A, et al. A systematic review of discontinued trials suggested that most reasons for recruitment failure were preventable. J Clin Epidemiol 2016 Dec;80:8-15. [doi: 10.1016/j.jclinepi.2016.07.016] [Medline: 27498376]
7. Colditz JB, Chu K, Emery SL, Larkin CR, James AE, Welling J, et al. Toward real-time infoveillance of Twitter health messages. Am J Public Health 2018 Aug;108(8):1009-1014. [doi: 10.2105/AJPH.2018.304497] [Medline: 29927648]
8. Eysenbach G. Infodemiology and infoveillance: Framework for an emerging set of public health informatics methods to analyze search, communication and publication behavior on the Internet. J Med Internet Res 2009 Mar 27;11(1):e11 [FREE Full text] [doi: 10.2196/jmir.1157] [Medline: 19329408]
9. Salathé M, Bengtsson L, Bodnar TJ, Brewer DD, Brownstein JS, Buckee C, et al. Digital epidemiology. PLoS Comput Biol 2012;8(7):e1002616 [FREE Full text] [doi: 10.1371/journal.pcbi.1002616] [Medline: 22844241]
10. Gelinas L, Pierce R, Winkler S, Cohen IG, Lynch HF, Bierer BE. Using social media as a research recruitment tool: Ethical issues and recommendations. Am J Bioeth 2017 Mar;17(3):3-14 [FREE Full text] [doi: 10.1080/15265161.2016.1276644] [Medline: 28207365]
11. Petronio S. Boundaries of Privacy: Dialectics of Disclosure. Albany, NY: State University of New York Press; 2002.
12. Marwick A, Boyd D. Networked privacy: How teenagers negotiate context in social media. New Media Soc 2014 Jul 21;16(7):1051-1067 [FREE Full text] [doi: 10.1177/1461444814543995]
13. Nissenbaum H. Privacy in Context: Technology, Policy, and the Integrity of Social Life. Stanford, CA: Stanford University Press; 2010.
14. Pew Research Center. Washington, DC: Pew Internet & American Life Project; 2019 Jun 12. Social media fact sheet URL: https://www.pewinternet.org/fact-sheet/social-media/ [accessed 2019-10-21]
15. Tsukayama H. The Washington Post. 2017 Nov 7. Twitter is officially doubling the character limit to 280 URL: https://www.washingtonpost.com/news/the-switch/wp/2017/11/07/twitter-is-officially-doubling-the-character-limit-to-280/ [accessed 2018-10-21]
16. Xu S, Markson C, Costello KL, Xing CY, Demissie K, Llanos AA. Leveraging social media to promote public health knowledge: Example of cancer awareness via Twitter. JMIR Public Health Surveill 2016;2(1):e17 [FREE Full text] [doi: 10.2196/publichealth.5205] [Medline: 27227152]
17. Sedrak MS, Cohen RB, Merchant RM, Schapira MM. Cancer communication in the social media age. JAMA Oncol 2016 Jun 01;2(6):822-823. [doi: 10.1001/jamaoncol.2015.5475] [Medline: 26940041]
18. Amon KL, Campbell AJ, Hawke C, Steinbeck K. Facebook as a recruitment tool for adolescent health research: A systematic review. Acad Pediatr 2014;14(5):439-447.e4. [doi: 10.1016/j.acap.2014.05.049] [Medline: 25169155]
19. Topolovec-Vranic J, Natarajan K. The use of social media in recruitment for medical research studies: A scoping review. J Med Internet Res 2016 Nov 07;18(11):e286 [FREE Full text] [doi: 10.2196/jmir.5698] [Medline: 27821383]
20. Golder S, Ahmed S, Norman G, Booth A. Attitudes toward the ethics of research using social media: A systematic review. J Med Internet Res 2017 Jun 06;19(6):e195. [doi: 10.2196/jmir.7082] [Medline: 28588006]
21. Bender JL, Cyr AB, Arbuckle L, Ferris LE. Ethics and privacy implications of using the Internet and social media to recruit participants for health research: A privacy-by-design framework for online recruitment. J Med Internet Res 2017 Apr 06;19(4):e104 [FREE Full text] [doi: 10.2196/jmir.7029] [Medline: 28385682]
22. Fiesler C, Proferes N. "Participant" perceptions of Twitter research ethics. Soc Media Soc 2018 Mar 10;4(1):205630511876336. [doi: 10.1177/2056305118763366]
23. Eysenbach G. Improving the quality of Web surveys: The Checklist for Reporting Results of Internet E-Surveys (CHERRIES). J Med Internet Res 2004 Sep 29;6(3):e34 [FREE Full text] [doi: 10.2196/jmir.6.3.e34] [Medline: 15471760]
24. National Institutes of Health. Washington, DC: National Institutes of Health; 2019 May 30. "Why should I participate in a clinical trial?" URL: https://www.nih.gov/health-information/nih-clinical-research-trials-you/why-should-i-participate-clinical-trial [accessed 2019-10-24]

25. Litman L, Robinson J, Abberbock T. TurkPrime.com: A versatile crowdsourcing data acquisition platform for the behavioral sciences. Behav Res Methods 2017 Apr;49(2):433-442 [FREE Full text] [doi: 10.3758/s13428-016-0727-z] [Medline: 27071389]
26. Teitcher JE, Bockting WO, Bauermeister JA, Hoefer CJ, Miner MH, Klitzman RL. Detecting, preventing, and responding to "fraudsters" in Internet research: Ethics and tradeoffs. J Law Med Ethics 2015;43(1):116-133 [FREE Full text] [doi: 10.1111/jlme.12200] [Medline: 25846043]
27. Cunningham JA, Godinho A, Hodgins DC. Pilot randomized controlled trial of an online intervention for problem gamblers. Addict Behav Rep 2019 Jun;9:100175 [FREE Full text] [doi: 10.1016/j.abrep.2019.100175] [Medline: 31193792]
28. Cunningham JA, Godinho A, Bertholet N. Outcomes of two randomized controlled trials, employing participants recruited through Mechanical Turk, of Internet interventions targeting unhealthy alcohol use. BMC Med Res Methodol 2019 Jun 14;19(1):124 [FREE Full text] [doi: 10.1186/s12874-019-0770-4] [Medline: 31200648]
29. Simone M, Geiser C, Lockhart G. The importance of face-to-face contact and reciprocal relationships and their associations with depressive symptoms and life satisfaction. Qual Life Res 2019 Nov;28(11):2909-2917. [doi: 10.1007/s11136-019-02232-7] [Medline: 31201729]
30. Jirsaraie RJ, Ranby KW, Albeck DS. Early life stress moderates the relationship between age and prosocial behaviors. Child Abuse Negl 2019 Aug;94:104029. [doi: 10.1016/j.chiabu.2019.104029] [Medline: 31207572]
31. Yap K, Grisham JR. Unpacking the construct of emotional attachment to objects and its association with hoarding symptoms. J Behav Addict 2019 Jun 01;8(2):249-258. [doi: 10.1556/2006.8.2019.15] [Medline: 31112034]
32. Bunge E, Cook HM, Bond M, Williamson RE, Cano M, Barrera AZ, et al. Comparing Amazon Mechanical Turk with unpaid Internet resources in online clinical trials. Internet Interv 2018 Jun;12:68-73 [FREE Full text] [doi: 10.1016/j.invent.2018.04.001] [Medline: 30135770]
33. Arditte KA, Çek D, Shaw AM, Timpano KR. The importance of assessing clinical phenomena in Mechanical Turk research. Psychol Assess 2016 Jun;28(6):684-691. [doi: 10.1037/pas0000217] [Medline: 26302105]
34. Liao Z, Elekdag-Turk S, Turk T, Grove J, Dalci O, Chen J, et al. Computational and clinical investigation on the role of mechanical vibration on orthodontic tooth movement. J Biomech 2017 Jul 26;60:57-64. [doi: 10.1016/j.jbiomech.2017.06.012] [Medline: 28743370]
35. Buccafusco C, Burns ZC, Fromer JC, Sprigman CJ. Experimental tests of intellectual property laws' creativity thresholds. Tex Law Rev 2014 Jun 01;93:1921-1980 [FREE Full text]
36. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap): A metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009 Apr;42(2):377-381 [FREE Full text] [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
37. Reuter K. figshare. 2019 Jun 25. Responses from Twitter users: Monitoring Twitter for clinical trial recruitment URL: https://figshare.com/articles/Responses_from_Twitter_users_Monitoring_Twitter_for_clinical_trial_recruitment/8308973 [accessed 2019-10-21]
38. Reuter K. figshare. 2019 Jun 25. Responses from TurkPrime workers: Monitoring Twitter for clinical trial recruitment URL: https://figshare.com/articles/Responses_from_TurkPrime_workers_Monitoring_Twitter_for_clinical_trial_recruitment/8309429 [accessed 2019-10-21]
39. Clyne W, Pezaro S, Deeny K, Kneafsey R. Using social media to generate and collect primary data: The #ShowsWorkplaceCompassion Twitter research campaign. JMIR Public Health Surveill 2018 Apr 23;4(2):e41 [FREE Full text] [doi: 10.2196/publichealth.7686] [Medline: 29685866]
40. Sinnenberg L, Buttenheim AM, Padrez K, Mancheno C, Ungar L, Merchant RM. Twitter as a tool for health research: A systematic review. Am J Public Health 2017 Jan;107(1):e1-e8. [doi: 10.2105/AJPH.2016.303512] [Medline: 27854532]
41. Kang R, Dabbish L, Fruchter N, Kiesler S. My data just goes everywhere: User mental models of the Internet and implications for privacy and security. In: Proceedings of the Eleventh USENIX Conference on Usable Privacy and Security. Berkeley, CA: USENIX Association; 2015 Jul Presented at: Eleventh USENIX Conference on Usable Privacy and Security; July 22-24, 2015; Ottawa, Canada p. 39-52 URL: https://dl.acm.org/citation.cfm?id=3235870
42. Park YJ. Digital literacy and privacy behavior online. Commun Res 2011 Aug 23;40(2):215-236. [doi: 10.1177/0093650211418338]
43. Bartsch M, Dienlin T. Control your Facebook: An analysis of online privacy literacy. Comput Human Behav 2016 Mar;56:147-154. [doi: 10.1016/j.chb.2015.11.022]
44. Chambers LA, Rueda S, Baker DN, Wilson MG, Deutsch R, Raeifar E, Stigma Review Team. Stigma, HIV and health: A qualitative synthesis. BMC Public Health 2015 Sep 03;15:848 [FREE Full text] [doi: 10.1186/s12889-015-2197-0] [Medline: 26334626]
45. Reuter K, Angyan P, Le N, MacLennan A, Cole S, Bluthenthal RN, et al. Monitoring Twitter conversations for targeted recruitment in cancer trials in Los Angeles County: Protocol for a mixed-methods pilot study. JMIR Res Protoc 2018 Sep 25;7(9):e177 [FREE Full text] [doi: 10.2196/resprot.9762] [Medline: 30274964]
46. McClure CC, Cataldi JR, O'Leary ST. Vaccine hesitancy: Where we are and where we are going. Clin Ther 2017 Aug;39(8):1550-1562. [doi: 10.1016/j.clinthera.2017.07.003] [Medline: 28774498]

47. Ross J, Irani I, Silberman MS, Zaldivar A, Tomlinson B. Who are the Turkers? Worker demographics in Amazon Mechanical Turk. In: Proceedings of ACM CHI 2010 Conference on Human Factors in Computing Systems. 2010 Presented at: ACM CHI 2010 Conference on Human Factors in Computing Systems; April 10-15, 2010; Atlanta, GA p. 2863-2872 URL: https://pdfs.semanticscholar.org/2261/5775c75ba7a227bce389532a6b2cae7c4ab7.pdf [doi: 10.1002/9781405165518.wbeos0994]

Abbreviations

AML: acute myeloid leukemia
CHERRIES: Checklist for Reporting Results of Internet E-Surveys
CIRB: Central Institutional Review Board
HPV: human papilloma virus
IRB: Institutional Review Board
REDCap: Research Electronic Data Capture

Edited by G Eysenbach; submitted 15.07.19; peer-reviewed by E Buchanan, S Venkatesan, S Guness; comments to author 11.08.19; revised version received 04.10.19; accepted 04.10.19; published 30.10.19.

Please cite as:
Reuter K, Zhu Y, Angyan P, Le N, Merchant AA, Zimmer M
Public Concern About Monitoring Twitter Users and Their Conversations to Recruit for Clinical Trials: Survey Study
J Med Internet Res 2019;21(10):e15455
URL: http://www.jmir.org/2019/10/e15455/
doi:10.2196/15455
PMID:31670698

Original Paper

Developing a Hypothetical Implementation Framework of Expectations for Monitoring Early Signs of Psychosis Relapse Using a Mobile App: Qualitative Study

Stephanie Allan[1], MA; Simon Bradstreet[1], PhD; Hamish Mcleod[1], PhD; John Farhall[2], PhD; Maria Lambrou[3], DClinPsy; John Gleeson[3], PhD; Andrea Clark[1], MSc; EMPOWER Group[1]; Andrew Gumley[1], PhD

[1]Mental Health and Wellbeing, Institute of Health & Wellbeing, University of Glasgow, Glasgow, United Kingdom

[2]Department of Psychology and Counselling, La Trobe University, Melbourne, Australia

[3]School of Behavioural and Health Sciences, Australian Catholic University, Melbourne, Australia

Corresponding Author:
Andrew Gumley, PhD
Mental Health and Wellbeing
Institute of Health & Wellbeing
University of Glasgow
Fleming Pavilion, West of Scotland Science Park
Glasgow, G20 0XA
United Kingdom
Phone: 44 01412113939
Email: andrew.gumley@glasgow.ac.uk

Abstract

Background: Relapse is a common experience for people diagnosed with psychosis, which is associated with increased service costs and profound personal and familial distress. EMPOWER (Early signs Monitoring to Prevent relapse in psychosis and prOmote Well-being, Engagement, and Recovery) is a peer worker–supported digital intervention that aims to enable service users to self-monitor their mental health with the aim of encouraging self-management and the shared use of personal data to promote relapse prevention. Digital interventions have not been widely used in relapse prevention and, therefore, little is currently known about their likely implementation—both within trials and beyond.

Objective: Seeking the perspectives of all relevant stakeholder groups is recommended in developing theories about implementation because this can reveal important group differences in understandings and assumptions about whether and for whom the intervention is expected to work. However, the majority of intervention implementation research has been retrospective. This study aimed to discover and theoretically frame implementation expectations in advance of testing and synthesize these data into a framework.

Methods: To develop a hypothetical implementation framework, 149 mental health professionals, carers, and people diagnosed with psychosis participated in 25 focus groups in both Australia and the United Kingdom. An interview schedule informed by the normalization process theory was used to explore stakeholders' expectations about the implementation of the EMPOWER intervention. Data were analyzed using thematic analysis and then theoretically framed using the Medical Research Council guidelines for understanding the implementation of complex interventions.

Results: All groups expected that EMPOWER could be successfully implemented if the intervention generated data that were meaningful to mental health staff, carers, and service users within their unique roles. However, there were key differences between staff, carers, and service users about what facilitators and barriers that stakeholders believe exist for intervention implementation in both the cluster randomized controlled trial stage and beyond. For example, service user expectations mostly clustered around subjective user experiences, whereas staff and carers spoke more about the impact upon staff interactions with service users.

Conclusions: A hypothetical implementation framework synthesized from stakeholder implementation expectations provides an opportunity to compare actual implementation data gathered during an ongoing clinical trial, giving valuable insights into the accuracy of these stakeholders' previous expectations. This is among the first studies to assess and record implementation expectations for a newly developed digital intervention for psychosis in advance of testing in a clinical trial.

Trial Registration: ISRCTN Registry ISRCTN99559262; http://www.isrctn.com/ISRCTN99559262

(J Med Internet Res 2019;21(10):e14366) doi:10.2196/14366

KEYWORDS

psychosis; self-management; implementation science

Introduction

Background

Relapse is common for many people diagnosed with schizophrenia [1]. Relapses are linked to increased disability from loss of important relationships and reduced education and employment opportunities [2]. One estimate suggests that psychotic relapse costs £10,950 (at 6 months) compared with £2532 for no relapse, with 75% of the difference in these costs coming from inpatient treatment [3]. In the United States, excess costs from relapse range from US $6033-$32,753 [4]. Commonly, relapses are preceded by so-called early warning signs (EWS) that reflect a combination of symptoms such as anxiety, depression, suspiciousness, and uniquely personal experiences. EWS-based prevention strategies assume that identifying relapse early enough enables preventative action and averts full relapse [5]. Guidelines for psychosis in both Scotland, the United Kingdom, [2] and Australia [6] recommend early signs-based strategies as crucial for relapse prevention in routine psychosis care.

Research into reliable and valid signs of relapse is essential for early intervention aimed at minimizing the harms associated with relapse [7]. A review [8] to determine the validity of early signs as predictors of relapse in people with nonaffective psychosis found that the sensitivity (correct relapse prediction by staff) ranged from 10% to 80% (median 61%), and specificity (nonrelapses correctly identified) ranged from 38% to 100% (median 81%). Therefore, existing systems used to identify EWS have an uncertain prognostic utility and may result in an unnecessary intervention that engenders fear of relapse in service users and carers [9]. Delayed help-seeking narrows the window for timely intervention [10] and can result in the use of coercive treatment measures that confirm negative expectations [11] and make disclosure of EWS more threatening for service users. Therefore, new interventions that address problems associated with help-seeking and disclosing EWS appear warranted [12].

Early Signs Monitoring to Prevent Relapse in Psychosis and Promote Well-Being, Engagement, and Recovery Description

One emerging application of technology in mental health care is remote self-monitoring [13]. Remote self-monitoring may improve upon traditional face-to-face monitoring by allowing more regular sampling of symptoms and, potentially, earlier detection of relapse signs. EMPOWER (Early signs Monitoring to Prevent relapse in psychosis and prOmote Well-being, Engagement, and Recovery; ISRCTN99559262) aims to develop and evaluate a mobile app for use with adults who experience psychosis. The app enables routine self-monitoring for a variety of different experiences, including psychosis (eg, hearing voices and suspicious thoughts), anxiety, mood, self-esteem, and interpersonal support. Furthermore, each time people complete an app questionnaire, they receive an *EMPOWER message*, which (depending on user input) provides links to further relevant information, practical advice, or helpful quotes. The EMPOWER algorithm aims to tailor these messages to individual changes in user well-being to promote a greater sense of control over mental health and to support self-management. EMPOWER participants will use the app for an initial 28-day baseline period to identify their typical variation in personal well-being. Significant changes from baseline will be then be triaged by a clinician. Peer support workers will be involved in setting up and personalizing the daily questionnaire, alongside regular fortnightly follow-up meetings where they will support service users in using the app.

Implementation of Digital Interventions

Digital interventions can help address clinical priorities in psychosis, such as increasing access to psychological interventions for symptoms such as paranoia [14]. However, many effective digital interventions have failed to generalize from clinical trials into clinical practice [15,16]. Owing to concerns about generalization beyond trial contexts, the UK Department of Health [17] encourages systematic implementation research to increase an understanding of how interventions are adopted or rejected. The effectiveness of interventions (including their success in reaching the target population) can be influenced by how an intervention interacts with the context in which it was implemented [18,19]. When appraising the results of a clinical trial, it can be challenging to know whether the intervention will generalize into *real-world* contexts of clinical practice. Process evaluations assess the implementation of interventions and help predict generalizability in different contexts. The Medical Research Council (MRC) framework for process evaluation [18] recommends clear descriptions of assumptions about how the intervention is expected to be implemented within a specific context. In addition, consulting multiple stakeholder groups is recommended because this can reveal across-group variance in understandings of what the intervention is and differences in assumptions about why and for whom the intervention is expected to work. Collecting data at different time points is also recommended to characterize changes in implementation factors such as participants' attitudes toward an intervention.

Typically, the majority of implementation research on engagement with interventions has been retrospective [20]. The MRC framework for process evaluations recommends that implementation research should proactively include key stakeholders because those expected to engage with an intervention are likely to have relevant experiential knowledge, which is useful in understanding the implementation process during a trial [18]. Qualitative research carried out *during* a trial (eg, asking service users about their experiences) can aid in understanding why an intervention might work and how context affects implementation [21]. However, *befor* e interacting with an intervention, stakeholders may have pre-existing expectations regarding implementation that will shape how they interact with

a planned intervention (hypothetical acceptability). Hypothetical acceptability is measured by key stakeholders' willingness to engage with a proposed intervention and in previous trials of digital interventions for severe mental health problems actual acceptability (assessed postintervention) is typically higher than hypothetical acceptability [22].

Theory in implementation science implies some predictive capacity [23]. Typically, implementation theory aims to create conceptual tools that enable researchers to describe, identify, and explain crucial elements of the implementation process and its outcomes [24]. Developing implementation theories in advance of empirical testing provides a framework for developing predictions about how interventions will interact with the context in which they are tested. Furthermore, completing this work allows researchers to make informed predictions about what implementation barriers that might be reasonably expected [25]. One such implementation theory, normalization process theory (NPT) [25] focuses on the work that groups and individuals do when interacting with an intervention and how they make sense of it, many intervention studies have successfully utilized NPT as a framework to guide research to more fully understand the implementation process [26]. Despite the recommended involvement of patients and members of the public within implementation research [27] and widespread assumptions that consultation work can help researchers anticipate stakeholders' needs, capacities, and priorities [28], the MRC guidelines on process evaluation [18] report substantial empirical uncertainty regarding the value of Patient and Public Involvement (PPI) work. However, stakeholders are likely to offer insights beyond the acceptability of digital interventions (eg, predicting intervention implementation barriers during testing) and arguably have a right to be involved in research, which impacts them. Adding the insight of carers, service users, and mental health staff should lead to a clearer understanding of barriers and facilitators to implementation.

To the best of our knowledge, only one other study has [29] explored staff, carers, and service users' perspectives of acceptability and implementation of a digital intervention for psychosis before engagement. Inclusion of these stakeholders enabled potentially diverse perspectives to be integrated into system design requirements for a mobile intervention for people who were considered to have treatment-resistant schizophrenia. Although this study is in a different population, the inclusion of multiple perspectives is a strength that could be applied to the prospective investigation of stakeholder engagement with digital interventions. In addition, there is little longitudinal research comparing stakeholder predictions pre-intervention with what happens when people interact with a digital intervention. Developing implementation theories for the EMPOWER intervention based on the expectations of staff, service users, and carers within a longitudinal process evaluation will allow for the assessment of the accuracy and the changing nature of these predictions over time, potentially highlighting the value of contextual knowledge that comes from consulting with stakeholders. We anticipate that developing an a priori implementation theory derived from stakeholder consultation will enhance implementation of the intervention in the context of a clinical trial and provide meaningful data to enable later generalization into clinical practice, a clear priority for services [15,17].

This study aimed to summarize the implementation expectations expressed within focus groups by mental health staff, carers, and service users in consultation work before a clinical trial to be able to compare these with the actual experiences of implementation observed within a feasibility study.

Methods

Design

This study forms part of the qualitative phase conducted before a cluster randomized controlled trial for the EMPOWER intervention (ISRCTN: 99559262). The methods are reported in line with the consolidated criteria for reporting qualitative research reporting recommendations for qualitative work [30]; a full checklist can be seen in Multimedia Appendix 1. Before the start of the study, ethical approvals were provided by West of Scotland REC (16/WS/0042) and Melbourne Health (REC/15/MH/344). Managerial approval was given by National Health Service Greater Glasgow and Clyde (NHSGG&C; GN14CP229) and North Western Mental Health Services (Project Number: 2015.286). The protocol is available in the National Institute of Health Research website [31]

Eligibility and Recruitment

All participants came from 1 health board area in the United Kingdom and 1 in Australia, where the intervention will be tested in a multisite clinical trial. Staff who support people with psychosis within Community Mental Health Services (CMHS) were invited to take part through initial researcher contact with clinical team leaders. Service user participants were invited to take part in focus groups through mental health staff and organizations providing support or representation to people with mental health difficulties. Service user participants were eligible if they were in contact with CMHS, had experienced a relapse within the previous 2 years, had received a diagnosis of Diagnostic and Statistical Manual of Mental Disorders-5 psychosis-related condition, and were able to provide informed consent. People who identified as carers for someone with psychosis were recruited from both mental health services and support organizations.

Focus Groups

Using focus groups rather than individual interviews enabled respondents to interact with and respond to the ideas and comments of other participants with whom they shared a role [32]. Focus groups were held in private rooms (of either CMHS or support organizations) and conducted by members of the research team using a topic guide. We did not collect demographic data beyond whether the participant was a carer, service user, or mental health staff. Following best practice guidelines [18], we used an explicit theoretical framework to guide our focus group schedule. An interview schedule informed by NPT [33] was developed to explore stakeholders' expectations. A copy of the topic guide for each of the stakeholder focus groups is provided in Multimedia Appendices 2-4.

A total of 25 focus groups were held across Melbourne and Glasgow from July 20, 2016, to September 9, 2017. Participants were 88 mental health staff, either working in the NHS in the United Kingdom (n=54, 9 focus groups) or NorthWestern Mental Health (public run) services in Australia (n=34, 4 focus groups). Focus group length ranged from 57 min to 2 hours and 9 min. A total of 21 service users were recruited from the United Kingdom (n=5, 3 focus groups) and Australia (n=16, 4 focus groups) and 40 carers from the United Kingdom (n=20, 2 focus groups) and Australia (n=20, 3 focus groups). Carers and service users received UK £20 or Aus $40 for participation. Staff received no cash reimbursement and participated during their usual working day. All participants gave written consent before taking part. All focus group facilitators (AG, SB, AC, ML, JG, JH, JF, and SA—a mix of genders) identified themselves as researchers to conduct the research and were transparent if they also held a clinical role. All participants received a presentation about the EMPOWER intervention. The focus groups were audio recorded and then transcribed verbatim. NVivo software (QSR International) was utilized to perform analysis.

Reflexivity

SA is a Doctor of Philosophy student investigating the implementation of digital interventions for psychosis. Facilitating focus groups was a task shared by all coauthors. Data analysis was primarily completed by SA, who has previously utilized qualitative methods. Supervision and code checking for all analysis (including discussions about saturation) were provided by AG and HM, both of whom are clinical psychologists. AG is chief investigator for the EMPOWER study and was responsible for the overall design and conduct of the research.

Data Analysis

The analysis comprised 2 stages. Thematic analysis is a qualitative method used to construct, analyze, and report on patterns within text data [34]. This is commonly utilized within qualitative aspects of process evaluations to identify key barriers and facilitators for implementation of a diverse range of digital interventions [35-37]. In stage 1, we performed an inductive thematic analysis [34] for each unique stakeholder group in turn. This was justified because in a pilot clinical trials such as EMPOWER, study evaluators are encouraged to use exploratory research to identify facilitators and barriers to interventions so that strategies can be put in place in time for an evaluation of effectiveness [18].

For stage 2, the MRC process evaluation framework [18] was identified as a suitable deductive coding framework [38] for placing the themes in an implementation theory context more relevant to the needs for a feasibility study where it may be too early to decide if normalization should be the goal. This was the rationale for moving away from our original plan (EMPOWER ISRCTN: 99559262) to use the NPT [39] framework for qualitative work. The MRC framework goes beyond barriers and facilitators to implementation and provides a taxonomy of implementation constructs. Expected barriers and facilitators (on their own) can be seen as singular aspects of a predicted overall process. However, during the analysis of focus group conversations, it was clear that barriers and facilitators were expected to *interact* together into an overall expected implementation process for EMPOWER. Therefore, we selected implementation constructs from the MRC process evaluation to structure our barriers and facilitators findings in a theoretically driven hypothetical implementation theory (presented as a deductive framework) for the EMPOWER trial:

- Reach (whether service users are expected to consent to take part)
- Fidelity (whether the intervention is expected to be used as described)
- Context (contextual factors expected to affect, or be affected by, the implementation process)
- Implementation (what successful implementation would look like in practice, beyond a trial)

Coding and analyzing the data within this framework resulted in the implementation issues highlighted during inductive analysis being more meaningfully constructed as implementation barriers and facilitators. Through our initial thematic analysis, we developed 16 themes (Table 1). The implementation diagram (Figure 1) represents implementation expectations for the EMPOWER intervention across staff, service users, and carers with facilitators (green) and barriers (red) within the implementation framework. The framework analysis was completed across all stakeholder groups simultaneously.

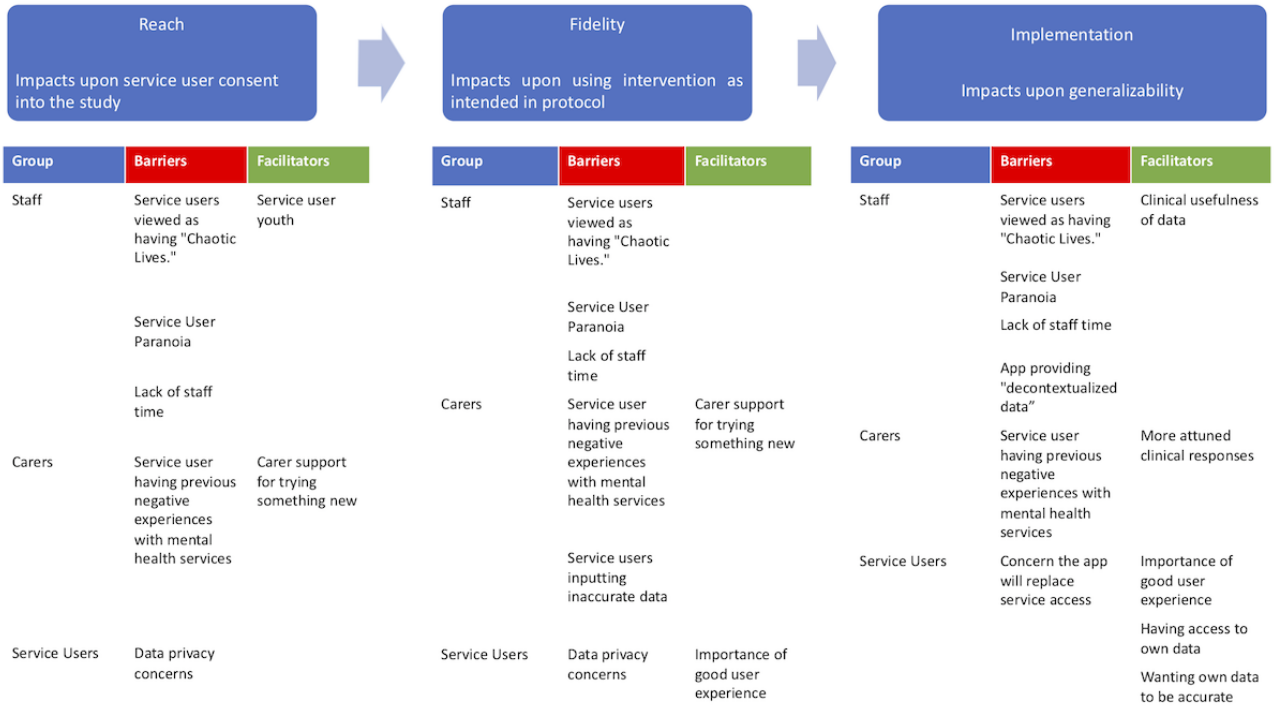
Both stages of qualitative analysis were completed by SA and triangulated through discussion with AG and HM. Resource limitations meant that strategies such as member checking (where participants check over themes proposed by the researcher as an interpretation validity check [40]) were not utilized. However, it has been highlighted that employing this technique may increase the validity of findings in qualitative research exploring user views of digital interventions in psychosis [41] and better ensure participant views have not been misrepresented.

Table 1. Themes from stage 1 analysis.

Stakeholder group	Expected implementation barriers	Expected implementation facilitators
Staff	Service users viewed as having *chaotic lives*	Service user youth
	Service user paranoia	Clinical usefulness of data
	Uncertainty about whether early warning signs data are useful in early intervention services	—[a]
	App providing *decontextualized data*	—
	Lack of staff time	—
Carers	Service user having previous negative experiences with mental health services	More attuned clinical responses
	Service users inputting inaccurate data	Carer support for trying something new
Service users	Data privacy concerns	Having access to own data
	Concern the app will replace service access	Wanting own data to be accurate
	—	Importance of good user experience

[a]Some cells are empty as there were fewer themes constructed.

Figure 1. A diagram of the hypothetical implementation framework.



Results

The first part of the Results section introduces the inductive thematic analysis (as shown in Table 1) and offers example quotes as an attempt to illustrate our analysis transparently.

Inductive Results

Mental Health Staff Implementation Expectations

Implementation Facilitators

Youth

Many staff predicted that young people (eg, those accessing early intervention services for psychosis) were more natural consumers of digital interventions. Staff perceived young service users as being both familiar with and highly able to use digital technology. Staff also expected that older service users would find the intervention harder to use and to be too burdensome for this reason. These assumptions appeared commonplace throughout discussions in both the United Kingdom and Australia:

> *I do think it's going to be a good thing in the long term, but there's going to be clients that don't fit into it now as well as. Because I think the next generation of people coming through are going to have been grown up with technology and are going to be okay with using it...* [Participant 8, Staff group 11, Australia]

Clinical Usefulness

Most staff appeared cautiously optimistic about the value of the data from the EMPOWER app and believed that it could be useful for their clinical practice by enabling staff to tune themselves into the changes in early signs and the broader context for these changes. In this particular illustrative quote, staff members highlighted how they expected EMPOWER data could draw their attention to patterns and links between stress and psychotic symptoms in the life of a service user:

> *You see where the stressors are, what times, what the patterns are, the patterns would be so clear.* [Participant 1, Staff group 2, United Kingdom]

Implementation Barriers

Service Users Viewed as Having Chaotic Lives

Staff reported that service users with a *chaotic* life would struggle to use the intervention. Staff viewed those individuals with chaotic lives as being the most vulnerable to relapse. *Chaotic lives* was a complex term referring to multiple factors including service users having difficulties with reflecting on their own experiences, having lack of insight, poor social or cognitive functioning, avoidance of services, or an inability to retain a mobile phone. These factors were considered in the context of the influence of a broader context of social deprivation or financial problems leading to users selling a provided mobile phone for cash:

> *It sounds like there'd be quite a specific group of patients that would benefit from this in terms of the people who are able to kind of reflect, who are you know, their lives aren't so chaotic that they can't keep hold of a mobile phone, you know, it doesn't end up somewhere else or in someone else's hands or whatever, and it's—I think it will be really useful for people who are functioning at that level and are able to reflect on things like that, but I guess it's—I suppose I'm just thinking it's a shame because it's often the people I suppose who I wonder might be at more risk of more kind of relapsing or being lost in the system somehow and becoming very unwell, are maybe already a bit too chaotic or functioning at too poor a level supposed to be able to make use of something as helpful potentially as this.* [Participant 1 Staff group 7, United Kingdom]

Service User Paranoia

Although the EMPOWER intervention was commonly described by staff to be an acceptable tool for managing relapse in at least some service users, they also perceived the intervention would not be acceptable to others. One common implementation barrier expected by staff was that service users with paranoid and/or delusional beliefs about technology would not engage with the intervention. This implementation expectation appeared grounded in expectations about how changing levels of paranoia will vary with technology affinity and competence. Conversations about service users who have technology-focused beliefs were frequent throughout staff focus groups and can be exemplified in the quote below where a staff member wonders aloud if EMPOWER would work for someone who already has such concerns about digital technology. Furthermore, this staff member highlighted that these beliefs could become more pronounced in the context of relapse:

> *I'm thinking about one of my service users in particular who, when he becomes unwell, his phone is actually part of his delusional belief system, and he becomes obsessive about certain part of it; so I'm wondering how that would work for him?* [Participant 4, Staff group 6, United Kingdom]

Uncertainty About Whether Early Warning Signs Data Are Useful in Early Intervention Services

Despite the optimistic expectation staff held about younger service users engaging with the intervention, staff from early intervention services discussed some different implementation barriers not present in other focus groups. For instance, the early stages of psychosis can be an uncertain time for clinicians because EWS of relapse might not be established yet. As illustrated below, a staff member from an early intervention service within the United Kingdom highlighted that the EMPOWER intervention might face a different implementation barrier because the data gathered via the app might have limited utility for staff in predicting relapse:

> *It's a trial but it is quite on the edge of relapse, which is risky. With our patient group, relapse signature is not that familiar because of early on. So, you've not got that history to learn from.* [Participant 2, Staff focus group 3, United Kingdom]

App Providing Decontextualized Data

Many staff expressed the concern that the quantitative self-reported data gathered from service users through their usage of the app lacked the context that comes from typical interactions staff have with service users. Overall, data alone were understood as being potentially unhelpful without the clinical experience of staff members to interpret these data. Staff valued their knowledge and relationship-based experiences of service users as a basis for making decisions concerning the risk of relapse. There was an additional concern that the quantity of data could also potentially block effective decision making. An example of this can be seen below where a staff member highlighted that if information from the EMPOWER app implies that a service user is relapsing, they would not feel comfortable acting on this information alone:

> *a bit of an overload of information perhaps if we're getting like you know three or whatever plus messages from the app a day and we'd need to do a management plan around...at presentation and a big limitation in that sort of context is that you don't...it's difficult to get a feel from the person about what is happening for the person...* [Participant 3]
>
> *missing out on the interpersonal context* [Participant 4, Staff group 12, Australia]

Lack of Staff Time

Staff were concerned that using the intervention in practice might be difficult. Working with people with a diagnosis of psychosis was described as a time-intensive part of their role.

Staff reported having many other competing demands on their time and limited resources to do their jobs. Staff frequently referred to a lack of capacity in the system and resource constraints. Several mental health staff even reported the lack of available resources within the mental health system and were concerned that digital technologies might one day replace their jobs. In the example below, the other participants in the focus group agree with participant 1, expressing concern about the potential lack of staff capacity for the implementation of EMPOWER:

> *It definitely makes sense, in that my only worry about it is that thinking about my caseload at the moment and I just don't know where we'd have the capacity to be working with it. [Sounds of Agreement from Other Participants]Particularly because it's psychosis and schizophrenia illness and how disabling that is...erm, to people.* [Participant 1, Staff group 2, United Kingdom]

Carers' Implementation Expectations

Implementation Facilitators

More Attuned Clinical Responses

Many carers expressed the view that routine monitoring and access to chart data could result in more attuned responses from mental health services because the data would indicate when support was needed. They believed that this would result in their loved one engaging with services when necessary, and services having a response that was experienced by their loved ones to be more relevant, timely, and acceptable. As demonstrated in the example below, carers state that they expect themselves to have a role in starting the help-seeking process:

> *if the chart was, you notice yourself it's is negative, they are definitely going down the tube, you will encourage them, if they don't see their doctor on a regular basis, that we should go and visit a doctor.* [Participant 2, Carer group 3, Australia]

Carer Support for Trying Something New

Aside from reporting implementation concerns for EMPOWER, carers also said it was essential to try out new interventions aimed at improving the lives of people with psychosis. Throughout all focus groups, it seemed clear that carers valued that clinical researchers were attempting to introduce innovation and were supportive of the role of research. Although carers were cautious about how successful their encouragement may be, they appeared keen to encourage ongoing usage of self-monitoring interventions by people who they support:

> *if we [as carers] had a good working understanding of it [EMPOWER] I'd find it easier to say to her "oh how are you getting on with the app?" and just encouraging her with it if she was happy to be encouraged, yeah. So, I think that'd be really good.* [Participant 5, Carer group 1, United Kingdom]

Implementation Barriers

Service User Having Previous Negative Experiences With Mental Health Services

However, similar to staff, carers frequently expressed that they expected the intervention to face multiple implementation barriers. Carers were nearly unanimous that the previous experiences of people with psychosis accessing services are likely to shape the reach of the intervention. This can be seen in the example below, where a carer predicts that her son is unlikely to use the EMPOWER intervention because of his previous autocratic experience dealing with mental health services. However, she remains cautiously optimistic about the implementation potential for the intervention of service users with different experiences:

> *I just...in my son's case, he wouldn't use it. He just wouldn't use it. And that's down to the experiences he's had with what he says is the mental health authorities. He's really...but for people who are open to it, it would be terrific. [murmur of agreement from other participants* [Participant 7, Carer group 5, Australia]

Service Users Inputting Inaccurate Data

Carers reported widespread concern that their loved ones may inaccurately input data. Throughout focus groups, this was understood as a function of concerns that their loved ones would downplay or minimize their experiences to avoid unwanted responses and interventions from services that they believe could result from accurate data input:

> *I suppose in some people if they are trying to be over positive and not give the truth.* [Participant 3, Carer group 1, United Kingdom]

Service Users' Implementation Expectations

Implementation Facilitators

Having Access to Own Data

Service users expected that having access to their data could be a useful source of learning about and becoming attuned to their well-being. Focus group discussions highlighted that psychotic experiences and general well-being are very changeable for service users. Data access appeared understood as a potential way to explore and learn about possible patterns, which might exist in these same well-being changes. In this particular example, a service user remarks that having data might encourage them to use the app because they feel that they are not currently aware of how their well-being fluctuates:

> *I would use them to see what's making me happy, what's doing my head in. How is my sleep schedule, am I getting ill. It's just understanding your own mind better than when you're doing it yourself. Because you're not really aware of all these things. You forget what you done yesterday.* [Participant 1, Service user group 1, United Kingdom]

Wanting Own Data to Be Accurate

Service users reported valuing having their data and expressed an awareness that for EMPOWER to work optimally, data entry will need to be accurate. In recognition of this, service user participants reflected on the importance of responding to the survey to the best of their ability. In the example below, a participant describes this being an implementation facilitator because inaccurate data would make the app data meaningless and would not confer any benefit:

> *don't lie to yourself because if you lying to the app the you are lying to yourself and basically you are not doing anyone any favours.* [Participant 2, Service user group 5, Australia]

Importance of App Providing a Good User Experience

Service users highlighted the importance of the app being appealing to use and the proposed message content being relevant and nonpatronizing. In the example below, a service user highlighted how they would feel infuriated if they were made to feel patronized. However, they stated that if they had control over what content they had to read, this would improve acceptability. Discussions such as this were commonplace and suggested that service users' perceptions of intervention content were a vital implementation expectation:

> *Participant 1: Yeah. There's a risk that it might be a wee bit patronising. Just a risk, I don't know. I know that me personally if I was feeling down in the dumps and I got a message saying "go for a walk"...* [laughs]
>
> *Researcher 1: "Pull your socks up."*
>
> *Participant 1: Yeah. It may infuriate me. But maybe if I had the option to read the message, I was choosing to read the message, it wouldn't be so annoying.* [Service user group 1, United Kingdom]

However, user experience conversations were not limited to intervention content. Discussions about the importance of how the app looks were common throughout focus groups. In the example below, a participant highlighted the importance of the intervention providing good experience through aesthetics. Therefore, the importance of user experience seems to envelop both intervention content as well as the package in which the intervention is delivered:

> *if it looks decent, if it doesn't look like a ten-year-old made it. Yeah. It has to be engaging and it look visually... that's pretty important to me. Not what I stand for, a ten-year-old* [Participant 3, Service user group 7, Australia]

Implementation Barriers

Data Privacy Concerns

Some service users stated that EMPOWER might be unacceptable to them because of expected paranoia. However, more common concerns were expressed regarding the privacy of data inputted into the app. The example below suggests that the service user is already concerned about threats to their privacy/autonomy and highlights that they are wary because their information will be sent to the treating team. Although this specific example highlights concern about information going to mental health staff, the focus group conversations also revealed concerns about other people, such as government employees or hackers, getting access to personal data. Therefore, this theme may reflect existing privacy concerns in the lives of service users. Although service users were generally accepting of the intervention regarding its role in supporting self-monitoring, they were cautious and guarded about being monitored by others, particularly mental health services:

> *Participant 3: We know that nothing is essentially private, well I happen to know that nothing that you tell any counsellor or social worker, nurse, therapist, anything, everything you tell them can be transferred even if it's just in the lounge in the kitchen during lunchtime "oh blah de blah de blah." We know they share information about us. We know they... um there is no privacy. Well I know it.*
>
> *Participant 1: Uh what was the question again?*
>
> *Researcher 1: It's really about the security arrangements and confidentiality with app as we have explained it, if there is any concerns or comments about that?*
>
> *Participant 3: Totally, it's going to be sending information to the treating team* [Service user group 7, Australia]

Concern App Will Replace Service Access

Service users throughout focus groups described accessing mental health services as a source of support in managing their well-being. The EMPOWER intervention was described as likely to encounter implementation barriers if the technical side of the intervention was perceived to be replacing *high-touch* human connection. In the example below, a service user participant highlighted that the digital intervention on its own would be a poor substitute for dealing with a person who knows them:

> *seems a poor substitute for seeing a person that knows you* [Participant 2, Service user group 1, United Kingdom]

Deductive Results

Barriers impacting upon reach (who consents to participate in the trial) are expected early in the implementation process. For example, carers expect that service users with previous negative experiences such as coercive treatment will be less likely to consent to the study (a reach barrier). Mental health staff expected that service users who have low general levels of functioning and/or high levels of paranoia would not consent or struggle to use the app if they do. However, mental health staff expected that younger service users would be more likely to be willing to participate in a digital intervention study because their generation are *digital natives*. Implementation issues that impact upon fidelity (such as service users inputting inaccurate data) are expected slightly later in the implementation process. However, even if the implementation is successful (with service users completing daily self-monitoring) and the data are perceived to be an accurate reflection of their mental state—problems in using EMPOWER data for relapse prevention are still expected. For example, staff predicted that

EMPOWER data will not be applicable within the context of early intervention services because EWS of relapse will still be unclear for people experiencing first episode psychosis (an implementation barrier). Barriers such as a lack of staff time were constructed as a predicted barrier across all levels (ie, expected to impact upon everything from service user consents into a feasibility study all the way up to generalizing into clinical practice if clinical outcomes in a definitive randomized controlled trial were favorable). The results of this deductive analysis can be seen in Figure 1.

Table 1 presents the themes as barriers and facilitators constructed during the inductive analysis.

Figure 1 presents the hypothetical implementation framework that scaffolds both barriers and facilitators themes that came up during focus group discussions. The diagram shows that throughout all stages, barriers and facilitators reach, fidelity, and implementation were constructed as coming from context.

Discussion

Principal Findings

This study is among the first to assess and record implementation expectations across mental health staff, carers, and service users for a newly developed digital intervention for psychosis in advance of testing in a clinical trial, building on previous multistakeholder work [29]. We have identified and theoretically framed the most common implementation expectations expressed by mental health staff, service users, and carers in advance of the EMPOWER clinical trial. Understanding the context behind empirical outcomes from novel digital mental health interventions is key in deciding if an intervention can be easily implemented within current practice [16] or will require significant resources and effort to do so [42]. Within a standard implementation science approach, context is defined as a shared environment, which can provide either barriers or facilitators for implementation [18]. However, within a complexity science–informed understanding, context is defined by an intervention interacting with multiple enacted environments of different social actors [43]. Although the MRC process evaluation framework provides a theoretical framework, creating the framework shown in Figure 1 means that it is more tailored to the clinical context of relapse management as reported by carers, mental health staff, and service users. Our findings provide a complexity science–informed account of how different stakeholders expect EMPOWER to interact within the multistakeholder actions that already occur during routine relapse prevention.

Key to the proposed framework (Figure 1) is a similarity between groups regarding expectations of what would constitute successful implementation. For successful implementation, it was agreed that EMPOWER must enable service user participants to self-monitor to a level of granularity that results in data allowing for visualization of potential personal indicators of relapse while also giving a comprehensive insight into overall service user mental health. Despite this implementation expectation appearing similar across groups, there were some role differences between staff, service users, and carers. The context of health care settings is constructed as being institutionalized [44] because behaviors by social actors are described in terms of the roles people are expected to act out. Our findings suggest that implementing the use of EMPOWER data in relapse prevention is only expected to be successful if the data are symbolically meaningful [15] to each stakeholder's role. For example, in the case of staff, this means having data that enables them to understand better how a participant feels and can help them differentiate EWS of relapse from a false alarm. For carers, useful data were constructed as staff becoming more attuned and being able to differentiate relapse signals from false alarms. Although both staff and carers emphasized data access as being an implementation facilitator that could improve service responses, service users were more curious about the impact of having access to a record of their self-reported day-to-day well-being. Previous qualitative research conducted with service users exploring potential [29,41,45] and actual [46] acceptability of digital self-management interventions for psychosis has reported that having access to personal data may have positive impacts such as enhancing self-management. However, this previous study also highlights more negative impacts reported by service users such as creating concerns about data privacy [41] and paranoia [46] and that using digital interventions may eventually lead to a reduction in mental health services [41]. Therefore, the mixed findings from our study appear mainly in line with previous research.

Similar to previous work exploring hypothetical implementation expectations held by staff, service users, and carers for a digital intervention for an online portal for schizophrenia [29], we found key differences in implementation expectations across staff, service users, and carers. Service user implementation expectations for both barriers and facilitators most frequently focused on individual experience. For example, the importance of EMPOWER providing a good user experience was highlighted as a key implementation facilitator throughout all stages of the implementation process and will be very important for sustained intervention use. User experience has been described as a neglected area within digital intervention research [47] and psychosis more specifically [48]. A recent study examining a mobile health platform for clinical monitoring in psychosis indicates that implementation was low because of the app frequently crashing [49], perhaps highlighting the importance of exploring user experience in implementation research. Carers (similar to findings from previous qualitative work [50]) and staff generally reflected how they foresee EMPOWER influencing service user interactions with staff. Furthermore, staff foreseeing digital interventions having an impact on staff roles and responsibilities is similar to previous qualitative research work conducted with mental health staff [29,51]. Carers expected that previous negative experiences of mental health care could act as a barrier toward initial engagement with the app. For carers, this expectation appeared to be related to a fear that EMPOWER would come to emulate existing dynamics within relapse prevention that can block timely communication of EWS. These findings are in line with previous research demonstrating that different stakeholders can hold different perspectives on digital mental health interventions [29,52,53] and suggest value in seeking out all relevant stakeholder perspectives.

This consultation work was helpful to the EMPOWER study because it highlighted key concerns of key stakeholders. For example, staff reporting a concern that app-generated data would be decontextualized data that may not be useful for clinical decision making. Going forward into the feasibility study, the role of a clinician in triaging data from the intervention to place app data within a meaningful context was emphasized to staff during recruitment.

Limitations

This study has several limitations. First, focus groups may result in some participants feeling reluctant to share their views fully. Second, the implementation barriers and facilitators highlighted in this paper were those that were most commonly discussed throughout the focus groups. However, the quantity of discussion of barriers and facilitators may not equal their importance or relevance [54]. Third, participants were given a *presentation* that covered the EMPOWER rationale and how the intervention works. Participants might have formed different expectations if they were presented with an *actual* prototype. A recent recommendation for undertaking complexity science–informed implementation research within health care services is to abandon attempts to simplify implementation research but rather explore implementation more inductively from multiple perspectives [55]. Therefore, there is a concern that adopting existing implementation taxonomy from the MRC process evaluation framework [18] within our analytic approach may have overly simplified construction of the hypothetical implementation framework. Moreover, following the NPT framework in designing research questions may have minimized the range of potential responses from participants. Finally, PPI can range from consultation to stakeholders having decision making over the aims and conduct of a study [56]. Therefore, these findings should be considered in light of them coming from consultation and not direct stakeholder involvement.

Conclusions

The field of digital self-monitoring interventions in psychosis is rapidly expanding [46,48], and there is a need to optimize interventions for implementation. One critical implementation-focused strategy is intervention co-design with stakeholders to develop digital psychosis interventions more suitable to the needs of end users [57,58]. After completion of the EMPOWER feasibility trial, we will utilize observations amassed during the trial to base comments on how stakeholder expectations identified from this analysis compare with actual trial implementation. This qualitative work done in advance of the EMPOWER trial provides insight into very early implementation expectations that form when people are first told about a digital intervention. These implementation expectations seem associated with the role that a person plays in managing a health problem (such as being a patient or a carer) as well as their previous experiences. Furthermore, these expectations extend across different levels of implementation [59], from early engagement to posttrial implementation—indicating that expectations are complex and wide ranging. Our results suggest that potential participants may quickly form implementation-related expectations about interventions and make predictions about how they (and others) will interact with the intervention. These findings indicate that potential participants do not arrive at interventions in a naïve state and may develop expectations and assumptions about new technology before they even use it for themselves.

Acknowledgments

The authors thank all the participants who agreed to take part in the study. This study was supported by NHS Research Scotland through the Chief Scientist Office and the Scottish Mental Health Research Network. This project was funded by the National Institute for Health Research Health Technology Assessment program (project number 13/154/04) and the National Heath Medical Research Council (APP1095879). It will be published in full in the Health Technology Assessment. The research was supported by NHS Research Scotland, through NHSGG&C and of the Mental Health Network. SA was funded by the Cremore Research Fellowship. The views and opinions expressed are those of the authors and do not necessarily reflect those of the Health Technology Assessment program, National Institute for Health Research, the NHS, The Department of Health, NorthWestern Mental Health Services, The Cremore Research Fund, or the National Health Medical Research Council. We also wish to thank Jennifer Hargrave who volunteered in cofacilitating 2 of the focus groups in Glasgow. The EMPOWER research group comprises John Ainsworth, Sandra Bucci, Shôn Lewis, Matthew Machin, Alison Yung (University of Manchester), Mario Alvarez, Sue Cotton, Reeva Lederman (Orygen, The National Centre of Excellence in Youth Mental Health, Melbourne, Australia), Max Birchwood, Swaran Singh, Andrew Thompson (University of Warwick), Andrew Briggs, Chris Williams (Glasgow Institute of Health and Wellbeing, University of Glasgow), Paul French (Manchester Metropolitan University), Graeme MacLennan (University of Aberdeen), Cathy Mihalopolous (Deakin University), John Norrie, Matthias Schwannauer (University of Edinburgh), Frank Reilly, Lesley Smith (Scottish Recovery Network), and Suresh Sundram (Monash University).

Conflicts of Interest

None declared.

Multimedia Appendix 1
Completed consolidated criteria for reporting qualitative research form.
[PDF File (Adobe PDF File), 490 KB - jmir_v21i10e14366_app1.pdf]

Multimedia Appendix 2
Staff focus group schedule.
[PDF File (Adobe PDF File), 244 KB - jmir_v21i10e14366_app2.pdf]

Multimedia Appendix 3
Service user focus group schedule.
[PDF File (Adobe PDF File), 280 KB - jmir_v21i10e14366_app3.pdf]

Multimedia Appendix 4
Carer focus group schedule.
[PDF File (Adobe PDF File), 226 KB - jmir_v21i10e14366_app4.pdf]

References

1. Robinson D, Woerner MG, Alvir JM, Bilder R, Goldman R, Geisler S, et al. Predictors of relapse following response from a first episode of schizophrenia or schizoaffective disorder. Arch Gen Psychiatry 1999 Mar;56(3):241-247. [doi: 10.1001/archpsyc.56.3.241] [Medline: 10078501]
2. Scottish Intercollegiate Guidelines Network (SIGN). 2013. Management of Schizophrenia: A National Clinical Guideline URL: https://www.sign.ac.uk/assets/sign131.pdf [accessed 2019-08-09]
3. SANE. 2006. Getting Well, Staying Well URL: http://www.sane.org.uk/uploads/think_twice.pdf [accessed 2019-08-07]
4. Pennington M, McCrone P. The cost of relapse in schizophrenia. Pharmacoeconomics 2017 Sep;35(9):921-936. [doi: 10.1007/s40273-017-0515-3] [Medline: 28534254]
5. Eisner E, Barrowclough C, Lobban F, Drake R. Qualitative investigation of targets for and barriers to interventions to prevent psychosis relapse. BMC Psychiatry 2014 Jul 16;14:201 [FREE Full text] [doi: 10.1186/1471-244X-14-201] [Medline: 25030092]
6. Galletly C, Castle D, Dark F, Humberstone V, Jablensky A, Killackey E, et al. Royal Australian and New Zealand college of psychiatrists clinical practice guidelines for the management of schizophrenia and related disorders. Aust N Z J Psychiatry 2016 May;50(5):410-472. [doi: 10.1177/0004867416641195] [Medline: 27106681]
7. Emsley R, Chiliza B, Asmal L, Harvey BH. The nature of relapse in schizophrenia. BMC Psychiatry 2013 Feb 8;13:50 [FREE Full text] [doi: 10.1186/1471-244X-13-50] [Medline: 23394123]
8. Eisner E, Drake R, Barrowclough C. Assessing early signs of relapse in psychosis: review and future directions. Clin Psychol Rev 2013 Jul;33(5):637-653. [doi: 10.1016/j.cpr.2013.04.001] [Medline: 23628908]
9. Gumley AI, MacBeth A, Reilly JD, O'Grady M, White RG, McLeod H, et al. Fear of recurrence: results of a randomized trial of relapse detection in schizophrenia. Br J Clin Psychol 2015 Mar;54(1):49-62. [doi: 10.1111/bjc.12060] [Medline: 25040487]
10. Birchwood M, Spencer E. Early intervention in psychotic relapse. Clin Psychol Rev 2001 Nov;21(8):1211-1226. [doi: 10.1016/S0272-7358(01)00105-2] [Medline: 11702513]
11. Gumley A, Karatzias A, Power K, Reilly J, McNay L, O'Grady M. Early intervention for relapse in schizophrenia: impact of cognitive behavioural therapy on negative beliefs about psychosis and self-esteem. Br J Clin Psychol 2006 Jun;45(Pt 2):247-260. [doi: 10.1348/014466505X49925] [Medline: 16719982]
12. NICE | The National Institute for Health and Care Excellence. 2014. Psychosis and Schizophrenia in Adults: Treatment and Management URL: https://www.nice.org.uk/guidance/cg178/evidence/full-guideline-490503565 [accessed 2019-08-09]
13. Rajagopalan A, Shah P, Zhang MW, Ho RC. Digital platforms in the assessment and monitoring of patients with bipolar disorder. Brain Sci 2017 Nov 12;7(11):pii: E150 [FREE Full text] [doi: 10.3390/brainsci7110150] [Medline: 29137156]
14. Garety PA, Ward T, Freeman D, Fowler D, Emsley R, Dunn G, et al. SlowMo, a digital therapy targeting reasoning in paranoia, versus treatment as usual in the treatment of people who fear harm from others: study protocol for a randomised controlled trial. Trials 2017 Nov 2;18(1):510 [FREE Full text] [doi: 10.1186/s13063-017-2242-7] [Medline: 29096681]
15. Greenhalgh T, Wherton J, Papoutsi C, Lynch J, Hughes G, A'Court C, et al. Beyond adoption: a new framework for theorizing and evaluating nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies. J Med Internet Res 2017 Nov 1;19(11):e367 [FREE Full text] [doi: 10.2196/jmir.8775] [Medline: 29092808]
16. Aref-Adib G, McCloud T, Ross J, O'Hanlon P, Appleton V, Rowe S, et al. Factors affecting implementation of digital health interventions for people with psychosis or bipolar disorder, and their family and friends: a systematic review. Lancet Psychiatry 2019 Mar;6(3):257-266. [doi: 10.1016/s2215-0366(18)30302-x]
17. Gov.UK. 2017. A Framework for Mental Health Research URL: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/665576/A_framework_for_mental_health_research.pdf [accessed 2019-08-07]
18. Moore GF, Audrey S, Barker M, Bond L, Bonell C, Hardeman W, et al. Process evaluation of complex interventions: Medical Research Council guidance. Br Med J 2015 Mar 19;350:h1258 [FREE Full text] [doi: 10.1136/bmj.h1258] [Medline: 25791983]

19. Petticrew M, Anderson L, Elder R, Grimshaw J, Hopkins D, Hahn R, et al. Complex interventions and their implications for systematic reviews: a pragmatic approach. Int J Nurs Stud 2015 Jul;52(7):1211-1216. [doi: 10.1016/j.ijnurstu.2015.01.004] [Medline: 25661527]
20. Sutcliffe K, Thomas J, Stokes G, Hinds K, Bangpan M. Intervention component analysis (ICA): a pragmatic approach for identifying the critical features of complex interventions. Syst Rev 2015 Oct 29;4:140 [FREE Full text] [doi: 10.1186/s13643-015-0126-z] [Medline: 26514644]
21. Bonell C, Fletcher A, Morton M, Lorenc T, Moore L. Realist randomised controlled trials: a new approach to evaluating complex public health interventions. Soc Sci Med 2012 Dec;75(12):2299-2306. [doi: 10.1016/j.socscimed.2012.08.032] [Medline: 22989491]
22. Berry N, Lobban F, Emsley R, Bucci S. Acceptability of interventions delivered online and through mobile phones for people who experience severe mental health problems: a systematic review. J Med Internet Res 2016 May 31;18(5):e121 [FREE Full text] [doi: 10.2196/jmir.5250] [Medline: 27245693]
23. Nilsen P. Making sense of implementation theories, models and frameworks. Implement Sci 2015 Apr 21;10:53 [FREE Full text] [doi: 10.1186/s13012-015-0242-0] [Medline: 25895742]
24. May C. Towards a general theory of implementation. Implement Sci 2013 Feb 13;8:18 [FREE Full text] [doi: 10.1186/1748-5908-8-18] [Medline: 23406398]
25. May CR, Johnson M, Finch T. Implementation, context and complexity. Implement Sci 2016 Oct 19;11(1):141 [FREE Full text] [doi: 10.1186/s13012-016-0506-3] [Medline: 27756414]
26. May CR, Cummings A, Girling M, Bracher M, Mair FS, May CM, et al. Using Normalization Process Theory in feasibility studies and process evaluations of complex healthcare interventions: a systematic review. Implement Sci 2018 Jun 7;13(1):80 [FREE Full text] [doi: 10.1186/s13012-018-0758-1] [Medline: 29879986]
27. Esmail L, Moore E, Rein A. Evaluating patient and stakeholder engagement in research: moving from theory to practice. J Comp Eff Res 2015 Mar;4(2):133-145. [doi: 10.2217/cer.14.79] [Medline: 25825842]
28. Oliver K, Kothari A, Mays N. The dark side of coproduction: do the costs outweigh the benefits for health research? Health Res Policy Syst 2019 Mar 28;17(1):33 [FREE Full text] [doi: 10.1186/s12961-019-0432-3] [Medline: 30922339]
29. Huerta-Ramos E, Escobar-Villegas MS, Rubinstein K, Unoka ZS, Grasa E, Hospedales M, M-RESIST Group, et al. Measuring users' receptivity toward an integral intervention model based on mhealth solutions for patients with treatment-resistant schizophrenia (m-RESIST): a qualitative study. JMIR Mhealth Uhealth 2016 Sep 28;4(3):e112 [FREE Full text] [doi: 10.2196/mhealth.5716] [Medline: 27682896]
30. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
31. National Institute for Health Research. EMPOWER: Early Signs Monitoring to Prevent Relapse and PrOmote Wellbeing, Engagement and Recovery URL: https://www.journalslibrary.nihr.ac.uk/programmes/hta/1315404#/ [accessed 2019-08-07]
32. Kruger RA, Casey MA. Focus Groups: A Practical Guide for Applied Research. London, UK: Sage Publications; 2014.
33. May CR, Mair F, Finch T, MacFarlane A, Dowrick C, Treweek S, et al. Development of a theory of implementation and integration: Normalization Process Theory. Implement Sci 2009 May 21;4:29 [FREE Full text] [doi: 10.1186/1748-5908-4-29] [Medline: 19460163]
34. Braun A, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
35. van Reijen M, Asscheman M, Vriend I, van Mechelen W, Verhagen E. Users' perspectives, opportunities, and barriers of the Strengthen Your Ankle app for evidence-based ankle sprain prevention: mixed-methods process evaluation for a randomized controlled trial. JMIR Rehabil Assist Technol 2018 Jul 6;5(2):e13 [FREE Full text] [doi: 10.2196/rehab.8638] [Medline: 29980497]
36. Maar MA, Yeates K, Perkins N, Boesch L, Hua-Stewart D, Liu P, et al. A framework for the study of complex mhealth interventions in diverse cultural settings. JMIR Mhealth Uhealth 2017 Apr 20;5(4):e47 [FREE Full text] [doi: 10.2196/mhealth.7044] [Medline: 28428165]
37. Morton K, Dennison L, Bradbury K, Band RJ, May C, Raftery J, et al. Qualitative process study to explore the perceived burdens and benefits of a digital intervention for self-managing high blood pressure in primary care in the UK. BMJ Open 2018 May 8;8(5):e020843 [FREE Full text] [doi: 10.1136/bmjopen-2017-020843] [Medline: 29739782]
38. Gale NK, Heath G, Cameron E, Rashid S, Redwood S. Using the framework method for the analysis of qualitative data in multi-disciplinary health research. BMC Med Res Methodol 2013 Sep 18;13:117 [FREE Full text] [doi: 10.1186/1471-2288-13-117] [Medline: 24047204]
39. Murray E, Treweek S, Pope C, MacFarlane A, Ballini L, Dowrick C, et al. Normalisation process theory: a framework for developing, evaluating and implementing complex interventions. BMC Med 2010 Oct 20;8:63 [FREE Full text] [doi: 10.1186/1741-7015-8-63] [Medline: 20961442]
40. Doyle S. Member checking with older women: a framework for negotiating meaning. Health Care Women Int 2007;28(10):888-908. [doi: 10.1080/07399330701615325] [Medline: 17987459]

41. Berry N, Lobban F, Bucci S. A qualitative exploration of service user views about using digital health interventions for self-management in severe mental health problems. BMC Psychiatry 2019 Jan 21;19(1):35 [FREE Full text] [doi: 10.1186/s12888-018-1979-1] [Medline: 30665384]
42. Torous J, Firth J. Bridging the dichotomy of actual versus aspirational digital health. World Psychiatry 2018 Feb;17(1):108-109 [FREE Full text] [doi: 10.1002/wps.20464] [Medline: 29352531]
43. Braithwaite J, Churruca K, Long JC, Ellis LA, Herkes J. When complexity science meets implementation science: a theoretical and empirical analysis of systems change. BMC Med 2018 Apr 30;16(1):63 [FREE Full text] [doi: 10.1186/s12916-018-1057-z] [Medline: 29706132]
44. Greenhalgh T, Shaw S, Wherton J, Vijayaraghavan S, Morris J, Bhattacharya S, et al. Real-world implementation of video outpatient consultations at macro, meso, and micro levels: mixed-method study. J Med Internet Res 2018 Apr 17;20(4):e150 [FREE Full text] [doi: 10.2196/jmir.9897] [Medline: 29625956]
45. Bucci S, Morris R, Berry K, Berry N, Haddock G, Barrowclough C, et al. Early psychosis service user views on digital technology: qualitative analysis. JMIR Ment Health 2018 Oct 31;5(4):e10091 [FREE Full text] [doi: 10.2196/10091] [Medline: 30381280]
46. Eisner E, Drake RJ, Berry N, Barrowclough C, Emsley R, Machin M, et al. Development and long-term acceptability of ExPRESS, a mobile phone app to monitor basic symptoms and early signs of psychosis relapse. JMIR Mhealth Uhealth 2019 Mar 29;7(3):e11568 [FREE Full text] [doi: 10.2196/11568] [Medline: 30924789]
47. Michie S, Yardley L, West R, Patrick K, Greaves F. Developing and evaluating digital interventions to promote behavior change in health and health care: recommendations resulting from an international workshop. J Med Internet Res 2017 Jun 29;19(6):e232 [FREE Full text] [doi: 10.2196/jmir.7126] [Medline: 28663162]
48. Bell IH, Lim MH, Rossell SL, Thomas N. Ecological momentary assessment and intervention in the treatment of psychotic disorders: a systematic review. Psychiatr Serv 2017 Nov 1;68(11):1172-1181. [doi: 10.1176/appi.ps.201600523] [Medline: 28669284]
49. Kumar D, Tully LM, Iosif AM, Zakskorn LN, Nye KE, Zia A, et al. A mobile health platform for clinical monitoring in early psychosis: implementation in community-based outpatient early psychosis care. JMIR Ment Health 2018 Feb 27;5(1):e15 [FREE Full text] [doi: 10.2196/mental.8551] [Medline: 29487044]
50. Lal S, Daniel W, Rivard L. Perspectives of family members on using technology in youth mental health care: a qualitative study. JMIR Ment Health 2017 Jun 23;4(2):e21 [FREE Full text] [doi: 10.2196/mental.7296] [Medline: 28645887]
51. Berry N, Bucci S, Lobban F. Use of the internet and mobile phones for self-management of severe mental health problems: qualitative study of staff views. JMIR Ment Health 2017 Nov 1;4(4):e52 [FREE Full text] [doi: 10.2196/mental.8311] [Medline: 29092809]
52. Carper MM, McHugh RK, Barlow DH. The dissemination of computer-based psychological treatment: a preliminary analysis of patient and clinician perceptions. Adm Policy Ment Health 2013 Mar;40(2):87-95. [doi: 10.1007/s10488-011-0377-5] [Medline: 22001968]
53. Strand M, Gammon D, Eng LS, Ruland C. Exploring working relationships in mental health care via an e-recovery portal: qualitative study on the experiences of service users and health providers. JMIR Ment Health 2017 Nov 14;4(4):e54 [FREE Full text] [doi: 10.2196/mental.8491] [Medline: 29138127]
54. Kooij L, Groen WG, van Harten WH. Barriers and facilitators affecting patient portal implementation from an organizational perspective: qualitative study. J Med Internet Res 2018 May 11;20(5):e183 [FREE Full text] [doi: 10.2196/jmir.8989] [Medline: 29752253]
55. Greenhalgh T, Papoutsi C. Studying complexity in health services research: desperately seeking an overdue paradigm shift. BMC Med 2018 Jun 20;16(1):95 [FREE Full text] [doi: 10.1186/s12916-018-1089-4] [Medline: 29921272]
56. Greenhalgh T, Hinton L, Finlay T, Macfarlane A, Fahy N, Clyde B, et al. Frameworks for supporting patient and public involvement in research: systematic review and co-design pilot. Health Expect 2019 Apr 22:- (epub ahead of print)(forthcoming). [doi: 10.1111/hex.12888] [Medline: 31012259]
57. Hardy A, London S, Ward T, Waller H, Kuipers E. The development of SlowMo: how inclusive, user-centred design research can improve psychological therapies for psychosis. JMIR Ment Health 2018:- (forthcoming)(forthcoming).
58. McClelland GT, Fitzgerald M. A participatory mobile application (app) development project with mental health service users and clinicians. Health Educ J 2018 Jun 5;77(7):815-827. [doi: 10.1177/0017896918773790]
59. McGinn CA, Grenier S, Duplantie J, Shaw N, Sicotte C, Mathieu L, et al. Comparison of user groups' perspectives of barriers and facilitators to implementing electronic health records: a systematic review. BMC Med 2011 Apr 28;9:46 [FREE Full text] [doi: 10.1186/1741-7015-9-46] [Medline: 21524315]

Abbreviations

CMHS: Community Mental Health Services
EMPOWER: Early signs Monitoring to Prevent relapse in psychosis and prOmote Well-being, Engagement, and Recovery
EWS: early warning signs

MRC: Medical Research Council
NHS: National Health Service
NHSGG&C: National Health Service Greater Glasgow and Clyde
NPT: normalization process theory
PPI: patient and public involvement

Edited by G Eysenbach; submitted 12.04.19; peer-reviewed by E Eisner, A Dodd; comments to author 08.05.19; revised version received 10.07.19; accepted 19.07.19; published 24.10.19.

Please cite as:
Allan S, Bradstreet S, Mcleod H, Farhall J, Lambrou M, Gleeson J, Clark A, EMPOWER Group, Gumley A
Developing a Hypothetical Implementation Framework of Expectations for Monitoring Early Signs of Psychosis Relapse Using a Mobile App: Qualitative Study
J Med Internet Res 2019;21(10):e14366
URL: https://www.jmir.org/2019/10/e14366
doi:10.2196/14366
PMID:31651400

Original Paper

Antismoking Advertisements and Price Promotions and Their Association With the Urge to Smoke and Purchases in a Virtual Convenience Store: Randomized Experiment

Lauren McCarl Dutra[1], MA, SCD; James Nonnemaker[2], PhD; Brian Bradfield[2], BSc; Nathaniel Taylor[2], MPH; Jamie Guillory[3], PhD; Ashley Feld[2], MPH; Annice Kim[2], PhD

[1]Center for Health Policy Science and Tobacco Research, RTI International, Berkeley, CA, United States

[2]Center for Health Policy Science and Tobacco Research, RTI International, Research Triangle Park, NC, United States

[3]Prime Affect Research, Dublin, Ireland

Corresponding Author:
Lauren McCarl Dutra, MA, SCD
Center for Health Policy Science and Tobacco Research
RTI International
2150 Shattuck Avenue, Suite 1200
Berkeley, CA, 94704
United States
Phone: 1 5106658297
Email: ldutra@rti.org

Abstract

Background: Point of sale (POS) advertising is associated with smoking initiation, current smoking, and relapse among former smokers. Price promotion bans and antismoking advertisements (ads) are 2 possible interventions for combating POS advertising.

Objective: The purpose of this analysis was to determine the influence of antismoking ads and promotions on urges to smoke and tobacco purchases.

Methods: This analysis examined exposure to graphic (graphic images depicting physical consequences of tobacco use) and supportive (pictures of and supportive messages from former smokers) antismoking ads and promotions in a virtual convenience store as predictors of urge to smoke and buying tobacco products among 1200 current cigarette smokers and 800 recent quitters recruited via a Web-based panel (analytical n=1970). We constructed linear regression models for urge to smoke and logistic regression models for the odds of purchasing tobacco products, stratified by smoking status.

Results: The only significant finding was a significant negative relationship between exposure to supportive antismoking ads and urge to smoke among current smokers (beta coefficient=−5.04, 95% CI −9.85 to −0.22; *P*=.04). There was no significant relationship between graphic antismoking ads and urge to smoke among current smokers (coefficient=−3.77, 95% CI −8.56 to 1.02; *P*=.12). Neither relationship was significant for recent quitters (graphic: coefficient=−3.42, 95% CI −8.65 to 1.81; *P*=.15 or supportive: coefficient=−3.82, 95% CI −8.99 to 1.36; *P*=.20). There were no significant differences in urge to smoke by exposure to promotions for current smokers (coefficient=−1.06, 95% CI −4.53 to 2.41; *P*=.55) or recent quitters (coefficient=1.76, 95% CI −2.07 to 5.59; *P*=.37). There were also no differences in tobacco purchases by exposure to graphic (current smokers: coefficient=0.93, 95% CI 0.67 to 1.29; *P*=.66 and recent quitters: coefficient=0.73, 95% CI 0.44 to 1.19; *P*=.20) or supportive (current smokers: coefficient=1.05, 95% CI 0.75 to 1.46; *P*=.78 and recent quitters: coefficient=0.73, 95% CI 0.45 to 1.18; *P*=.20) antismoking ads or price promotions (current smokers: coefficient=1.09, 95% CI 0.86 to 1.38; *P*=.49 and recent quitters: coefficient=0.90, 95% CI 0.62 to 1.31; *P*=.60).

Conclusions: The results of this analysis support future research on the ability of supportive antismoking ads to reduce urges to smoke among current cigarette smokers. Research on urges to smoke has important tobacco control implications, given the relationship between urge to smoke and smoking cigarettes, time to next smoke, and amount smoked.

(J Med Internet Res 2019;21(10):e14143) doi:10.2196/14143

KEYWORDS

cigarette smoking; advertisement; craving; tobacco products; commerce; consumer behavior

Introduction

Background

Tobacco advertising promotes tobacco use, which results in 480,000 deaths each year in the United States [1]. In 2015, tobacco companies spent approximately US $8.5 billion, or 95% of their advertising budget, on the tobacco retail environment, otherwise known as the point of sale (POS) [2]. POS marketing influences susceptibility to smoking among youth [3,4] and quitting behavior among adults [5,6].

Moreover, 2 common and effective POS marketing techniques are tobacco displays (ie, large, colorful displays of tobacco products often referred to as power walls [7]) and price promotions, such as coupons and multipack discounts [8]. Tobacco displays and other forms of protobacco advertising at the POS have been associated with cravings to smoke among current and former tobacco users [9]. Tobacco displays are also associated with susceptibility to smoking among youth, fewer successful quit attempts among adults, and unplanned purchases of tobacco products among tobacco users [3,5,6,10-12].

Promotions are used by tobacco companies to offset price increases caused by tobacco control policies, such as taxes [13]. Promotions have been associated with current smoking among youth [4] and purchasing larger quantities of cigarettes per store visit among adult smokers [14]. Researchers found that New York State counties with a greater number of retail cigarette promotions between 2004 and 2008 also had a higher youth smoking prevalence [4]. Similarly, another study of combustible tobacco users living in the rural United States found that those who used promotions purchased more cigarettes [14].

Furthermore, 1 potential method of counteracting the effects of tobacco displays and promotions is antitobacco ads [9]. These ads, particularly those with emotional components (such as personal stories and graphic images), have been associated with higher odds of quitting smoking among US adult smokers [15]. In 2009, New York City posted graphic antismoking warning signs in tobacco retail stores [9]. After the warning signs were posted, visitors to New York City retail stores who viewed the warning signs (and protobacco advertising) were significantly more likely to report that the signs made them think about the health risks of smoking or quitting smoking compared with those who visited the stores before the warning signs were posted [9]. Similarly, graphic warning labels on cigarette packs have been associated with lower cravings to smoke [16].

However, 1 study suggests that antismoking ads (referred to hereafter as antismoking ads) may not be effective [17,18]. In a virtual shopping experiment [17], researchers exposed adult current cigarette smokers and recent quitters to either a closed tobacco display with no advertising or a closed display with a graphic antismoking ad. The researchers did not find any differences in urge to smoke or purchase attempts based on exposure to the antismoking ads. The US Food and Drug Administration implemented the first national POS antismoking media campaign, Every Try Counts [18], in 2018. The campaign targets adult cigarette smokers, particularly those who are trying to quit smoking cigarettes despite multiple failed quit attempts. The campaign involves placing supportive antismoking ads that depict former smokers who appear to be in good health, and the ads contain prosmoking cessation messages such as If at first you don't succeed, try, try again. These messages are designed to promote quit attempts among adult cigarette smokers in convenience stores around the United States. Evaluation of the campaign is currently underway.

Another approach to countering protobacco advertising at the POS is banning price promotions. These bans prevent retailers from discounting tobacco products (such as buy 1 get 1 free or US $1 off) and, therefore, have the potential to influence consumer tobacco purchases [19]. Several counties and US states have enacted price promotion bans. As of March 2018, 13 local governments in Massachusetts (Chelsea and Winthrop), Michigan (East Lansing), Minnesota (North Branch and Wyoming), New York (New York City), Rhode Island (Providence and Central Falls), Texas (Rockport and Magnolia), and Washington (Cheney, Spokane, and Millwood) had passed regulations to counteract price promotions on tobacco products. These areas passed minimum prices on cigarette packs, prohibited or restricted the ability of retailers to redeem coupons or use price-reduction promotions (eg, multipack discounts), and set a minimum pack size and prices for tobacco products other than cigarettes (eg, cigarillos and cigars) [20].

Objectives

To further understand the potential effects of antismoking ads, price promotions, and their combined effect on urges to smoke and tobacco purchases, we used RTI iShoppe (iShoppe), a virtual convenience store developed by RTI International, to conduct an experiment that used a virtual convenience store to vary these aspects of the retail environment. Virtual stores, which simulate a retail shopping experience, are helpful for evaluating the effects of new or potential tobacco control approaches. We created different versions of convenience stores in iShoppe to test the effects of antismoking ads (graphic ads, supportive ads, or no ads) and price promotions (present vs absent) on urges to smoke and tobacco purchases among current and former cigarette smokers. Graphic ads contain graphic depictions of the physical consequences of tobacco use. Supportive ads include pictures of and supportive messages from former smokers who appear healthy. This research has the potential to contribute to the existing evidence base for the use of antismoking ads and price promotion bans (as well as a combination of the 2) as tobacco control measures at the POS.

Given the evidence establishing a relationship between warning images and cravings to smoke [16], we hypothesized that participants exposed to antismoking ads would report lower urges to smoke and purchase fewer tobacco products in iShoppe. We also hypothesized that participants exposed to price promotions would report greater urges to smoke and be more likely to purchase tobacco products than those who were not exposed to promotions. In addition, given the differences in responses to advertising by smoking status [21-23], we hypothesized that smoking status would serve as an effect modifier of the relationship between the tobacco control measures and outcomes examined in this analysis.

Methods

Participants

We used Lightspeed's Web-based survey panel to recruit a national convenience sample of 1200 adult current cigarette smokers and 800 recent quitters. First, potential participants completed a screening survey to ensure that they met the inclusion criteria. Current cigarette smokers or recent quitters older than 18 years were eligible to participate. Current smokers were participants who had smoked at least 100 cigarettes in their lifetimes and who reported that they currently smoked every day or some days. Recent quitters were participants who had smoked at least 100 cigarettes in their lifetimes, were now smoking not at all, and reported that they had quit smoking within the past year.

Study Procedures

iShoppe is a 3-dimensional (3D) virtual environment based on an off-the-shelf model of a convenience store that was extensively customized using the Unity 3D interaction gaming software. Since the original version of the virtual store [24], which was based on feedback from focus groups, the store has been updated many times.

For this study, Lightspeed provided participants with a link to access the store. If the participant already had the Unity 3D player installed, the store loaded once the participant clicked on the link. If it was not installed, participants were provided with instructions on how to download it. The first screen displayed contained a list of instructions on how to use the store, including the keystrokes used to explore the store. Each participant was provided a budget of US $15 or US $20 for the experiment. In areas with higher tobacco product prices, a US $20 budget was given to ensure all participants could purchase tobacco products. Participants were informed of their budget and instructed to purchase whatever they wanted to purchase in the store (within the budget). Participants were provided instructions for completing their purchases and were given 10 min to complete the shopping task. Further information about iShoppe is available from previous publications [4,17,24-26]. After completing the shopping task, participants were routed to a survey that measured urge to smoke, recall of products and ads in the virtual store, usual tobacco purchasing behavior, and tobacco use.

Experimental Design

The study used a partially crossed 3 (antismoking ad type: graphic, supportive, or none) by 2 (antismoking ad placement: purchasable ad space only or purchasable ad space plus high-visibility ad space) by 2 (promotions: absent or present) design. Study conditions with no antismoking ads did not include variation by ad placement, making the design partially crossed. The experimental design contained 10 conditions (Figures 1-4, Multimedia Appendices 1 and 2). Approximately 200 participants were assigned to each condition (ie, 1200 current smokers and 800 recent quitters). However, this analysis focuses on the effects of antismoking ads and price promotions, but not placement of antismoking ads, because we were most interested in the main effects of antismoking ads and promotions. As a result, when we conducted the analysis, we collapsed conditions 3 and 5 (graphic ads and promotions banned [Figure 3]), 7 and 9 (supportive ads and promotions banned), 4 and 6 (graphic ads and promotions present), and 8 and 10 (supportive ads [Figure 4] and promotions present). Therefore, this analysis contains 2 variables (ad placement not included) and 6 conditions (Multimedia Appendix 3).

Figure 1. Condition 1: Antismoking advertisements are absent, and price promotions are banned.

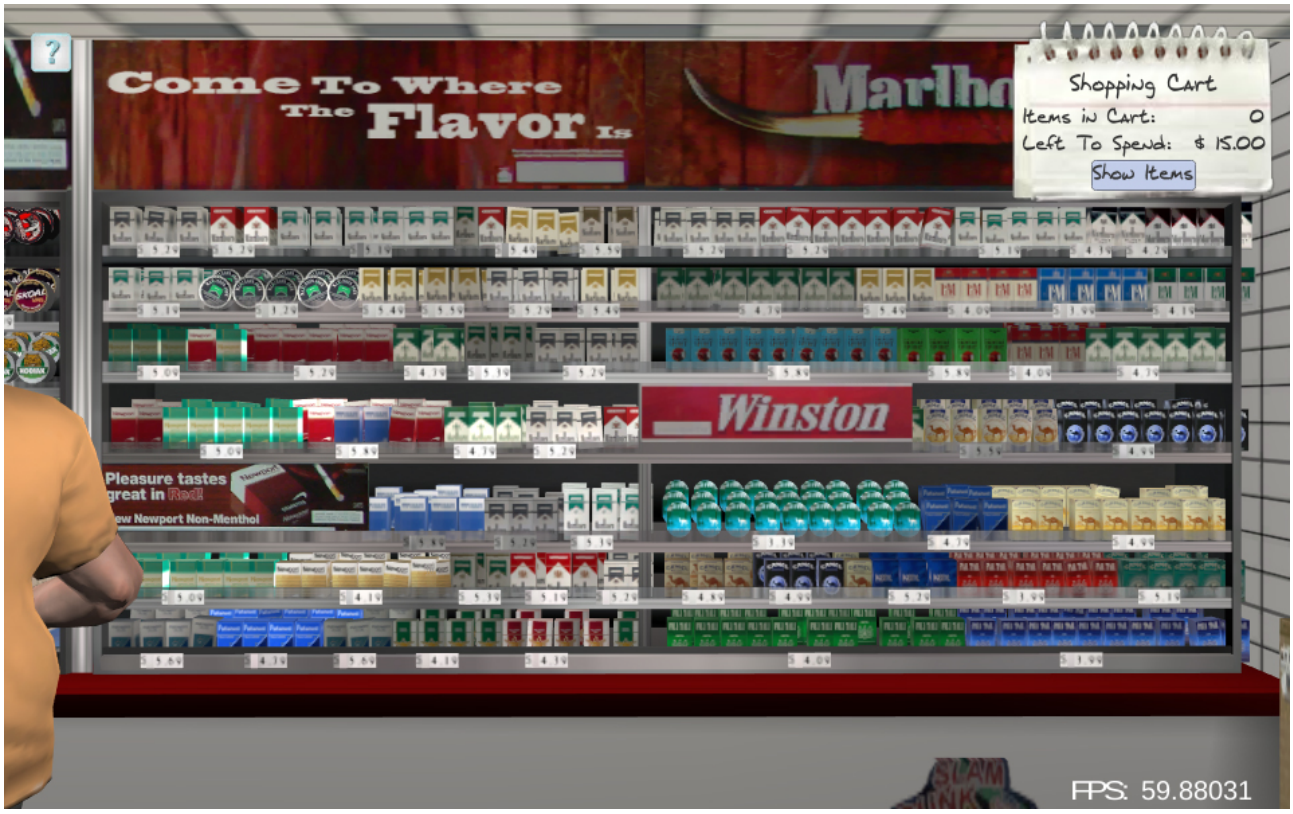


Figure 2. Condition 4: Graphic advertisements (ads) are present in purchasable ad space (eg, interior and exterior windows, gas pump topper) and price promotions are present (text on poster on pillar reads "SPECIAL OFFER: Buy two packs, get one free").

Figure 3. Condition 5: Graphic advertisements (ads) are present in purchasable (eg, interior and exterior windows, gas pump topper) and high visibility (eg, by checkout counter, hanging from the ceiling between aisles) ad space, and price promotions are banned.

Figure 4. Condition 7: Supportive advertisements (ads) are present in purchasable ad space (eg, interior and exterior windows, gas pump topper), and price promotions are banned.

Variables

Outcome Variables

Urge to smoke was assessed immediately after completing the virtual store shopping task, on a scale from 0 to 100, with 0 indicating no urge and 100 indicating the strongest urge I have ever experienced. Purchasing 1 or more tobacco products of any kind (by clicking to purchase) was also an outcome variable. The tobacco products available for purchase in the virtual store were cigarettes, cigars (including cigarillos and little cigars), smokeless tobacco, and electronic cigarettes (e-cigarettes).

Independent Variable

The exposure variables were antismoking ad condition (graphic, supportive, or neither, the last of which is the reference category) and price promotion condition (present or absent, the latter of which is the reference category).

Antismoking Advertisement Condition

We used antismoking ads from Centers for Disease Control and Prevention's (CDC) *Tips from Former Smokers* (*Tips)* national campaign. We chose these ads because they included graphic and supportive messages but were otherwise similar. We removed the CDC logo and placed national Quitline information (You can quit. Call 1-800-QUIT-NOW) at the bottom of all ads for consistency. Antismoking ads were placed in locations outside (gas pump toppers, walls, sandwich boards, ice chests, windows, and store entrance doors) and inside (interior windows and drink coolers) of the store that are typically available for purchase in most convenience stores. In the high-visibility condition, antismoking ads were also placed above the checkout counter (overhang) and hung from the ceiling in each center aisle.

Price Promotion Condition

The price promotion condition included special prices on specific items (eg, special price: US $0.50 off each pack) and discounts for buying more than 1 of the same product type (ie, multipack discounts). Promotions were placed on the tobacco display and ad posters for leading brands of specific tobacco products, including cigarettes, little cigars or cigarillos, smokeless tobacco, and e-cigarettes. The prices on the ads were customized based on each participant's geographic location to reflect the tobacco product pricing and regulations in each state. In all conditions, the store contained a visible tobacco display behind the checkout counter and ads for tobacco products.

Covariates

Covariates included age (18-24 [reference category], 25-34, 35-54, and ≥55 years), gender (male [reference category] or female), race and ethnicity (non-Hispanic white [reference category], non-Hispanic black, Hispanic, or non-Hispanic, other race), education (less than high school, high school graduate, some college, or a college degree or greater [reference category]), and the frequency at which the participant visited convenience stores (coded as a range from hardly ever [1] to daily [5]). We also examined the number of days smoked in the past 30 days among current smokers.

Statistical Analysis

We excluded participants with missing values for our outcome variables of interest, resulting in an analytic sample of 1970 participants (98.50% of the original sample of 2000 participants).

Descriptive Statistics

We tested for an imbalance in demographic covariates across conditions, which can sometimes occur even with randomization. Demographic covariates were included in multivariable models if they varied by experimental condition at the $P<.10$ level [27].

Bivariate Statistics

Next, we conducted bivariate analyses to understand the relationship between antismoking ads, price promotions, and the outcome variables using *t* tests for urge to smoke by price promotion; a 1-way analysis of variance for urge to smoke by antismoking ad condition; and chi-square analyses for tobacco purchases by promotion and antismoking ad condition.

Interactions

Then, to assess the interaction effects of the 2 independent variables, we tested interaction terms for antismoking ad by price promotions. Interaction terms that were significant at the $P<.10$ level were included in final regression models.

Covariate-Adjusted Regression Models

Finally, we constructed linear regression models for urge to smoke and logistic regression models for tobacco purchases. As we hypothesized that there would be differences in reactions to the store environment based on smoking status, we stratified all models by smoking status. We included all covariates in regression models that varied at the $P<.10$ level, with the exception of days smoked in the past 30, which we did not include because the variable was only relevant for current smokers (not for recent quitters).

Results

Descriptive Statistics

More than half of participants were current smokers (1177/1970, 59.75%); 40.25% (793/1970) of the participants were recent quitters (Table 1). The largest age group included participants aged from 35 to 54 years (636/1970, 32.28%), and the majority of the sample was female (1349/1970, 68.48%). The sample was primarily non-Hispanic white (1581/1970, 77.04%), and almost half of the sample (859/1970, 43.62%) had a college degree or greater education. On average, current smokers in the sample reported smoking on 25.6 (SD 8.2) of the past 30 days.

Table 1. Demographics characteristics of participants in the RTI iShoppe antismoking ad and price promotion study (n=1970).

Characteristics	Values
Smoking status, n (%)	
Current smoker	1177 (59.75)
Recent quitter	793 (40.25)
Age (years), n (%)	
18-24	271 (13.76)
25-34	519 (26.35)
35-54	636 (32.28)
≥55	544 (27.61)
Female, n (%)	1349 (68.46)
Race, n (%)	
Non-Hispanic white	1518 (77.04)
Non-Hispanic black	111 (5.64)
Hispanic	189 (9.58)
Non-Hispanic other	152 (7.74)
Education, n (%)	
Less than high school	42 (2.14)
High school graduate or General Educational Development degree	428 (21.71)
Some college	640 (32.54)
College graduate and beyond	859 (43.62)
Days smoked in the past 30 (current smokers only), mean (SD)	25.6 (8.2)

As expected, because of randomization, there were few differences in demographic characteristics between conditions. A greater proportion of participants in the antismoking ad condition (513/1579, 32.49%) reported attending some college compared with those in the no antismoking ad condition (111/391, 28.5%; *P*=.049). The antismoking ad condition also had fewer white participants (1146/1579, 76.2%) than the no antismoking ad condition (314/391, 80.2%; *P*=.08). Owing to these differences, race and ethnicity and education were included in all multivariable models.

Bivariate Statistics

Tobacco Purchases

There was no difference in tobacco purchases by antismoking ad condition (*P*=.78) or price promotion condition (*P*=.87; Table 2).

Urge to Smoke

Bivariate analyses of the outcome variables by antismoking ad (Table 2) revealed a lower urge to smoke among participants who viewed antismoking ads (at the *P*<.10 level [*P*=.08]). However, there was no difference in urge to smoke by presence of a price promotion (*P*=.58).

Table 2. In-store behaviors and self-reported urge to smoke cigarettes by smoking status among participants in the RTI iShoppe antismoking ad and price promotion study (n=1970).

Variable	Antismoking ad condition				Price promotions		
	Graphic (n=789)	Supportive (n=790)	Neither (n=391)	*P* value	Banned (n=981)	Present (n=989)	*P* value
Purchased any tobacco, n (%)	338 (42.8)	345 (43.7)	176 (45.01)	.78	426 (43.4)	433 (43.8)	.87
Urge to smoke (1-100), mean (SE)	42.1 (1.2)	40.3 (1.1)	44.8 (1.6)	.08	41.8 (1.1)	42.1 (1.0)	.58

Interactions

None of the interaction terms tested was significant at the *P*<.10 level for urge to smoke (current smokers: *P*=.25 for graphic ads by promotions and *P*=.91 for supportive ads by promotions; recent quitters: *P*=.93 for graphic ads by promotions and *P*=.21 for supportive ads by promotions), and therefore, the models of urge to smoke did not contain interaction terms. In the models of tobacco purchases, *P* values were .08 (graphic) and .27 (supportive) for interactions for current smokers and .34 (graphic) and .65 (supportive) among recent quitters. We did not include interaction terms for current smokers because the

results of the model remained virtually the same (graphic antismoking ad: odds ratio [OR]=0.09, 95% CI 0.45 to 1.76; supportive antismoking ad: OR=0.65, 95% CI 0.33 to 1.30; promotion: OR=0.99, 95% CI 0.46 to 2.12).

Covariate-Adjusted Regression Models

Urge to Smoke

Adjusting for covariates, current smokers exposed to supportive antismoking ads (coefficient=–5.04, 95% CI −9.85 to −0.22; *P*=.04; Table 3) reported significantly lower urges to smoke after visiting iShoppe than current smokers who were not exposed to the antismoking ads. However, there was no difference in urge to smoke based on exposure to graphic ads (coefficient=–3.77, 95% CI −8.56 to 1.02; *P*=.12) among current smokers (Table 3).

Among recent quitters, there was no difference in urge to smoke between those who did and did not see antismoking ads (graphic: coefficient=−3.42, 95% CI −8.65 to 1.81; *P*=.20 and supportive: coefficient=−3.82, 95% CI −8.99 to 1.36; *P*=.15).

There were no significant differences in urge to smoke between participants exposed to price promotions (vs not exposed) among current smokers (coefficient=−1.06, 95% CI −4.53 to 2.41; *P*=.55) or recent quitters (coefficient=1.76, 95% CI −2.07 to 5.59; *P*=.37).

Table 3. Multivariable regression models of urge to smoke (linear) and tobacco purchases (logistic) regressed on antismoking ad condition and presence of price promotions in the RTI iShoppe virtual convenience store (n=1970).

Exposure variable name, categories	Urge to smoke		Bought tobacco	
	Current smokers, beta coefficient (95% CI)	Recent quitters, beta coefficient (95% CI)	Current smokers, odds ratio (95% CI)	Recent quitters, odds ratio (95% CI)
Antismoking ad condition				
Graphic	−3.77 (−8.56 to 1.02)	−3.42 (−8.65 to 1.81)	0.93 (0.67 to 1.29)	0.73 (0.44 to 1.19)
Supportive	−5.04 (−9.85 to −0.22)	−3.82 (−8.99 to 1.36)	1.05 (0.75 to 1.46)	0.73 (0.45 to 1.18)
Neither	Reference	Reference	Reference	Reference
Price promotions				
Banned	Reference	Reference	Reference	Reference
Present	−1.06 (−4.53 to 2.41)	1.76 (−2.07 to 5.59)	1.09 (0.86 to 1.38)	0.90 (0.62 to 1.31)
Race				
Non-Hispanic white	Reference	Reference	Reference	Reference
Non-Hispanic black	5.23 (−2.37 to 12.83)	−5.25 (−13.58 to 3.09)	0.58 (0.35 to 0.96)	0.80 (0.33 to 1.94)
Hispanic	11.80 (5.71 to 17.88)	2.22 (−4.25 to 8.68)	0.96 (0.63 to 1.45)	1.16 (0.64 to 2.10)
Non-Hispanic other	6.83 (0.18 to 13.48)	6.59 (−0.57 to 13.75)	0.60 (0.38 to 0.93)	1.88 (1.03 to 3.46)
Education				
Less than high school	13.40 (1.93 to 24.88)	−7.03 (−22.16 to 8.10)	0.52 (0.24 to 1.11)	2.22 (0.66 to 7.47)
High school or General Educational Development degree	3.15 (−1.31 to 7.61)	0.62 (−4.66 to 5.90)	1.12 (0.82 to 1.52)	1.33 (0.79 to 2.25)
Some college	−5.65 (−9.83 to −1.48)	−2.64 (−6.99 to 1.71)	0.89 (0.67 to 1.19)	1.60 (1.05 to 2.45)
College degree or greater education	Reference	Reference	Reference	Reference

Tobacco Product Purchases

There were no significant differences in the odds of purchasing tobacco products between participants exposed to graphic (current smokers: OR 0.93, 95% CI 0.67 to 1.29; *P*=.66 and recent quitters: OR 0.73, 95% CI 0.44 to 1.19; *P*=.20) or supportive (current smokers: OR 1.05, 95% CI 0.75 to 1.46; *P*=.78 and recent quitters: OR 0.73, 95% CI 0.43 to 1.18; *P*=.20) ads versus those not exposed to antismoking ads (Table 3). Similarly, there was no difference in tobacco purchases by presence of price promotions (current smokers: OR 1.09, 95% CI 0.86 to 1.38; *P*=.49 and recent quitters: OR 0.90, 95% CI 0.62 to 1.31; *P*=.60; Table 3).

Discussion

Principal Findings

In a virtual convenience store shopping experiment conducted with adults, this analysis found a lower urge to smoke among current smokers who viewed supportive antismoking ads than those who did not view antismoking ads. On average, participants exposed to supportive antismoking ads reported urges to smoke that were 5 points lower (or 5% lower as the scale was 0-100) than those of participants not exposed to antismoking ads. This finding has important implications for the potential benefits of supportive antismoking ads at the POS. Although the direction of the effect was consistent with the results for supportive ads among current smokers, we found no

effect of supportive antismoking ads for recent quitters (quit smoking in the past year) or for graphic antismoking ads regardless of smoking status. We also tested whether purchasing 1 or more tobacco products while in the virtual store varied by exposure to antismoking ads or price promotions among current smokers and recent quitters. However, none of these relationships were significant.

Importance of Findings

Our finding of a lower urge to smoke among current smokers exposed to supportive ads suggests that supportive ads may be a method of decreasing urges to smoke among cigarette smokers. Urges to smoke are an important determinant of actual smoking behavior among current smokers [28] and self-efficacy to quit smoking among smokers attempting to quit smoking [29].

Comparison With Prior Work

Regarding our specific finding of the relationship between urge to smoke and supportive antiads, we were unable to find any previous study that examined this relationship in a virtual convenience store.

In terms of graphic ads, 1 other study [17] has examined variation in urge to smoke by exposure to graphic ads in a virtual convenience store and found results consistent to ours. Using a different convenience sample of adult current smokers and recent quitters than this analysis, Kim et al [17] examined urge to smoke while exposing participants to an open (standard) or enclosed tobacco display in iShoppe. Kim et al [17] found no difference in urge to smoke between participants exposed to a graphic antismoking ad versus no antismoking ad on the enclosed display.

Regarding price promotions, we were able to locate 1 study that examined the relationship between exposure to POS advertising, including tobacco promotions such as special prices, multipack discounts, or free gifts with purchase of cigarettes, and cravings to smoke [30]. Siahpush et al [6] found a positive relationship between promotions and cravings to smoke that approached significance (P=.06). Our findings for price promotions did not approach this level of significance; however, our experimental design and analysis also included antismoking ads, which Siahpush et al [6] did not include.

Unfortunately, we were unable to find any existing research that examined the relationship between exposure to price promotions and tobacco purchases. The only related study examined the relationship between the use of price promotions when making a tobacco purchase and the size of the purchase [14]. Despite the lack of research on the impact of price promotions on purchases in the United States, bans on price promotions have been cited as one of the most effective tobacco control efforts in Europe [19].

Limitations

Several limitations apply to this analysis. Owing to our use of a convenience sample, the results of this analysis may not generalize to all current cigarette smokers and recent quitters in the United States. However, specific aspects of our experimental design, such as randomization and adjusting for covariates, created comparable experimental conditions and, thereby, minimized threats to internal validity. After participants were divided into multiple experimental conditions, our sample sizes were rather small. However, collapsing the purchasable ad space only and high-visibility conditions improved sample sizes somewhat. In addition, it remains unclear whether the results of this experiment can be generalized to behavior in brick-and-mortar convenience stores or to Web purchases. However, existing research has found similarities between virtual store purchases and real-life purchases [31-34]. In addition, because we only used antismoking ads from *Tips*, the results of this analysis may not generalize to all antismoking ads. Finally, it is possible that some participants in the virtual store experiment did not notice the ads. However, because the study focused on antismoking ad exposure, we could not also adjust for awareness of the antismoking ads because awareness only varies among participants who have viewed the ads.

Conclusions

This analysis supports the potential utility of future research on the ability of supportive antismoking ads to combat urges to smoke among current cigarette smokers. Given that urge to smoke is an important predictor of smoking behavior, research should continue to explore the utility of antismoking ads as a method of influencing tobacco purchases at the POS. Given the existing research suggesting that the context and type of antismoking ads in stores can affect attention to and reactions to antismoking ads [35,36], if antismoking ads are used, the choice and placement of antismoking ads should be carefully considered when using these ads as a tobacco control intervention.

Acknowledgments

This research was funded by the US National Institutes of Health, National Institute on Drug Abuse, and the US Food and Drug Administration Center for Tobacco Products (grant P50DA036128).

Conflicts of Interest

None declared.

Multimedia Appendix 1
Table depicting the partial cross-over randomized design that resulted in 10 experimental groups.
[PDF File (Adobe PDF File), 105 KB - jmir_v21i10e14143_app1.pdf]

Multimedia Appendix 2
Images of the experimental conditions.
[PDF File (Adobe PDF File), 1224 KB - jmir_v21i10e14143_app2.pdf]

Multimedia Appendix 3
Flow diagram demonstrating how participants were assigned to 10 experimental conditions and then condensed into 6 conditions for the analysis.
[PDF File (Adobe PDF File), 234 KB - jmir_v21i10e14143_app3.pdf]

References

1. Surgeon General, US. The Health Consequences of Smoking--50 Years of Progress: A Report of the Surgeon General. Atlanta, GA: US Department of Health and Human Services; 2014.
2. Federal Trade Commission. 2017. Cigarette Report for 2015 URL: https://www.ftc.gov/system/files/documents/reports/federal-trade-commission-cigarette-report-2015-federal-trade-commission-smokeless-tobacco-report/2015_cigarette_report.pdf [accessed 2019-03-27]
3. Mackintosh AM, Moodie C, Hastings G. The association between point-of-sale displays and youth smoking susceptibility. Nicotine Tob Res 2012 May;14(5):616-620. [doi: 10.1093/ntr/ntr185] [Medline: 21908461]
4. Kim AE, Loomis BR, Busey AH, Farrelly MC, Willett JG, Juster HR. Influence of retail cigarette advertising, price promotions, and retailer compliance on youth smoking-related attitudes and behaviors. J Public Health Manag Pract 2013;19(6):E1-E9. [doi: 10.1097/PHH.0b013e3182980c47] [Medline: 23715218]
5. Germain D, McCarthy M, Wakefield M. Smoker sensitivity to retail tobacco displays and quitting: a cohort study. Addiction 2010 Jan;105(1):159-163. [doi: 10.1111/j.1360-0443.2009.02714.x] [Medline: 19804457]
6. Siahpush M, Shaikh RA, Smith D, Hyland A, Cummings KM, Kessler AS, et al. The association of exposure to point-of-sale tobacco marketing with quit attempt and quit success: results from a prospective study of smokers in the United States. Int J Environ Res Public Health 2016 Feb 6;13(2):203 [FREE Full text] [doi: 10.3390/ijerph13020203] [Medline: 26861379]
7. Martino SC, Setodji CM, Dunbar MS, Shadel WG. Increased attention to the tobacco power wall predicts increased smoking risk among adolescents. Addict Behav 2019 Jan;88:1-5 [FREE Full text] [doi: 10.1016/j.addbeh.2018.07.024] [Medline: 30098502]
8. Xu X, Wang X, Caraballo RS. Is every smoker interested in price promotions? An evaluation of price-related discounts by cigarette brands. J Public Health Manag Pract 2016;22(1):20-28 [FREE Full text] [doi: 10.1097/PHH.0000000000000223] [Medline: 26598952]
9. Coady MH, Chan CA, Auer K, Farley SM, Kilgore EA, Kansagra SM. Awareness and impact of New York City's graphic point-of-sale tobacco health warning signs. Tob Control 2013 May;22(e1):e51-e56. [doi: 10.1136/tobaccocontrol-2011-050261] [Medline: 22730446]
10. Clattenburg EJ, Elf JL, Apelberg BJ. Unplanned cigarette purchases and tobacco point of sale advertising: a potential barrier to smoking cessation. Tob Control 2013 Nov;22(6):376-381. [doi: 10.1136/tobaccocontrol-2012-050427] [Medline: 23138525]
11. Carter OB, Mills BW, Donovan RJ. The effect of retail cigarette pack displays on unplanned purchases: results from immediate postpurchase interviews. Tob Control 2009 Jun;18(3):218-221. [doi: 10.1136/tc.2008.027870] [Medline: 19264731]
12. Wakefield M, Germain D, Henriksen L. The effect of retail cigarette pack displays on impulse purchase. Addiction 2008 Feb;103(2):322-328. [doi: 10.1111/j.1360-0443.2007.02062.x] [Medline: 18042190]
13. Ross H, Tesche J, Vellios N. Undermining government tax policies: common legal strategies employed by the tobacco industry in response to tobacco tax increases. Prev Med 2017 Dec;105S:S19-S22 [FREE Full text] [doi: 10.1016/j.ypmed.2017.06.012] [Medline: 28625418]
14. Doogan NJ, Cooper S, Quisenberry AJ, Brasky TM, Browning CR, Klein EG, et al. The role of travel distance and price promotions in tobacco product purchase quantity. Health Place 2018 May;51:151-157 [FREE Full text] [doi: 10.1016/j.healthplace.2018.03.009] [Medline: 29625358]
15. Durkin SJ, Biener L, Wakefield MA. Effects of different types of antismoking ads on reducing disparities in smoking cessation among socioeconomic subgroups. Am J Public Health 2009 Dec;99(12):2217-2223 [FREE Full text] [doi: 10.2105/AJPH.2009.161638] [Medline: 19833980]
16. Lin PN, Zimmermann MH, Manderski MT, Schmelzer AC, Steinberg MB. Evaluation of graphic cigarette warning images on cravings to smoke. J Smok Cessat 2012 Feb 21;6(2):85-88. [doi: 10.1375/jsc.6.2.85]
17. Kim AE, Nonnemaker JM, Loomis BR, Shafer PR, Shaikh A, Hill E, et al. Influence of point-of-sale tobacco displays and graphic health warning signs on adults: evidence from a virtual store experimental study. Am J Public Health 2014 May;104(5):888-895. [doi: 10.2105/AJPH.2013.301723] [Medline: 24625149]

18. Food and Drug Administration. 2018. Every Try Counts: Campaign Overview URL: https://www.fda.gov/downloads/TobaccoProducts/PublicHealthEducation/PublicEducationCampaigns/EveryTryCountsCampaign/UCM588441.pdf [accessed 2019-03-14]
19. Schaap MM, Kunst AE, Leinsalu M, Regidor E, Ekholm O, Dzurova D, et al. Effect of nationwide tobacco control policies on smoking cessation in high and low educated groups in 18 European countries. Tob Control 2008 Aug;17(4):248-255 [FREE Full text] [doi: 10.1136/tc.2007.024265] [Medline: 18483129]
20. Truth Initiative. 2017. How Local Governments Are Countering Industry Efforts to Sell Cheap Tobacco URL: https://truthinitiative.org/news/how-local-governments-are-countering-industry-efforts-sell-cheap-tobacco [accessed 2019-02-12]
21. Yang D. The communication effects of audience situation and message framing on smoking cessation. Soc Sci Res Netw 2013;2(4):252-264. [doi: 10.2139/ssrn.2368006]
22. Schneider TR, Salovey P, Pallonen U, Mundorf N, Smith NF, Steward WT. Visual and auditory message framing effects on tobacco smoking. J Appl Social Pyschol 2001 Apr;31(4):667-682. [doi: 10.1111/j.1559-1816.2001.tb01407.x]
23. Sutfin EL, Szykman LR, Moore MC. Adolescents' responses to anti-tobacco advertising: exploring the role of adolescents' smoking status and advertisement theme. J Health Commun 2008;13(5):480-500. [doi: 10.1080/10810730802198961] [Medline: 18661389]
24. Kim AE, Nonnemaker JM, Loomis BR, Baig A, Hill E, Holloway JW, et al. Influence of tobacco displays and ads on youth: a virtual store experiment. Pediatrics 2013 Jan;131(1):e88-e95 [FREE Full text] [doi: 10.1542/peds.2012-0197] [Medline: 23209108]
25. Guillory J, Kim AE, Nonnemaker JM, Bradfield B, Taylor NH, Dutra L, et al. Effect of menthol cigarette and other menthol tobacco product bans on tobacco purchases in the RTI iShoppe virtual convenience store. Tob Control 2019 Jun 5:- (epub ahead of print). [doi: 10.1136/tobaccocontrol-2019-054997] [Medline: 31167902]
26. Dutra LM, Nonnemaker J, Taylor N, Kim AE. Deception and shopping behavior among current cigarette smokers: a web-based, randomized virtual shopping experiment. JMIR Res Protoc 2018 Jun 29;7(6):e10468 [FREE Full text] [doi: 10.2196/10468] [Medline: 29959114]
27. Bursac Z, Gauss CH, Williams DK, Hosmer DW. Purposeful selection of variables in logistic regression. Source Code Biol Med 2008 Dec 16;3:17 [FREE Full text] [doi: 10.1186/1751-0473-3-17] [Medline: 19087314]
28. Shiffman S, Dunbar M, Kirchner T, Li X, Tindle H, Anderson S, et al. Smoker reactivity to cues: effects on craving and on smoking behavior. J Abnorm Psychol 2013 Feb;122(1):264-280 [FREE Full text] [doi: 10.1037/a0028339] [Medline: 22708884]
29. Niaura R, Shadel WG, Britt DM, Abrams DB. Response to social stress, urge to smoke, and smoking cessation. Addict Behav 2002;27(2):241-250. [doi: 10.1016/s0306-4603(00)00180-5] [Medline: 11817765]
30. Siahpush M, Shaikh RA, Cummings KM, Hyland A, Dodd M, Carlson L, et al. The association of point-of-sale cigarette marketing with cravings to smoke: results from a cross-sectional population-based study. Tob Control 2016 Jul;25(4):402-405 [FREE Full text] [doi: 10.1136/tobaccocontrol-2015-052253] [Medline: 26024797]
31. Aubin G, Béliveau MF, Klinger E. An exploration of the ecological validity of the virtual action planning-supermarket (VAP-S) with people with schizophrenia. Neuropsychol Rehabil 2018 Jul;28(5):689-708. [doi: 10.1080/09602011.2015.1074083] [Medline: 26317526]
32. Desmet P, Bordenave R, Traynor J. Differences in purchasing behaviour between physical and virtual laboratory stores. Recherche et Applications en Marketing 2013 Jul 9;28(2):70-85. [doi: 10.1177/2051570713487473]
33. van Herpen E, van den Broek E, van Trijp HC, Yu T. Library Wageningen University & Research. 2014. Using a Virtual Grocery Store to Simulate Shopping Behaviour URL: https://library.wur.nl/WebQuery/wurpubs/490235 [accessed 2019-09-02]
34. Waterlander WE, Jiang Y, Steenhuis IH, Mhurchu CN. Using a 3D virtual supermarket to measure food purchase behavior: a validation study. J Med Internet Res 2015 Apr 28;17(4):e107 [FREE Full text] [doi: 10.2196/jmir.3774] [Medline: 25921185]
35. Dutra LM, Nonnemaker J, Guillory J, Bradfield B, Taylor N, Kim A. Smokers' attention to point-of-sale anti-smoking ads: an eye-tracking study. Tob Regul Sci 2018 Jan 1;4(1):631-643 [FREE Full text] [doi: 10.18001/TRS.4.1.9]
36. Kim A, Nonnemaker J, Guillory J, Shafer P, Parvanta S, Holloway J, et al. Antismoking ads at the point of sale: the influence of ad type and context on ad reactions. J Health Commun 2017 Jun;22(6):477-487. [doi: 10.1080/10810730.2017.1311970] [Medline: 28441097]

Abbreviations

3D: 3-dimensional
CDC: Centers for Disease Control and Prevention
e-cigarettes: electronic cigarettes
OR: odds ratio
POS: point of sale

Edited by G Eysenbach; submitted 30.04.19; peer-reviewed by Z Zhang, A Al-Zalabani; comments to author 17.07.19; revised version received 07.08.19; accepted 19.08.19; published 23.10.19.

Please cite as:
Dutra LM, Nonnemaker J, Bradfield B, Taylor N, Guillory J, Feld A, Kim A
Antismoking Advertisements and Price Promotions and Their Association With the Urge to Smoke and Purchases in a Virtual Convenience Store: Randomized Experiment
J Med Internet Res 2019;21(10):e14143
URL: https://www.jmir.org/2019/10/e14143
doi:10.2196/14143
PMID:31647468

Original Paper

A Virtual Counseling Application Using Artificial Intelligence for Communication Skills Training in Nursing Education: Development Study

Shefaly Shorey[1], PhD; Emily Ang[1], DNurs; John Yap[2], MA; Esperanza Debby Ng[1], BA; Siew Tiang Lau[1], PhD; Chee Kong Chui[3], PhD

[1]Alice Lee Centre for Nursing Studies, National University of Singapore, Singapore, Singapore

[2]Information Techonology, National University of Singapore, Singapore, Singapore

[3]Department of Mechanical Engineering, National University of Singapore, Singapore, Singapore

Corresponding Author:
Shefaly Shorey, PhD
Alice Lee Centre for Nursing Studies
National University of Singapore
Clinical Research Centre 10 Medical Drive
Singapore, 117597
Singapore
Phone: 65 66011294
Email: nurssh@nus.edu.sg

Related Article:

This is a corrected version. See correction statement: http://www.jmir.org/2019/11/e17064/

Abstract

Background: The ability of nursing undergraduates to communicate effectively with health care providers, patients, and their family members is crucial to their nursing professions as these can affect patient outcomes. However, the traditional use of didactic lectures for communication skills training is ineffective, and the use of standardized patients is not time- or cost-effective. Given the abilities of virtual patients (VPs) to simulate interactive and authentic clinical scenarios in secured environments with unlimited training attempts, a virtual counseling application is an ideal platform for nursing students to hone their communication skills before their clinical postings.

Objective: The aim of this study was to develop and test the use of VPs to better prepare nursing undergraduates for communicating with real-life patients, their family members, and other health care professionals during their clinical postings.

Methods: The stages of the creation of VPs included preparation, design, and development, followed by a testing phase before the official implementation. An initial voice chatbot was trained using a natural language processing engine, Google Cloud's Dialogflow, and was later visualized into a three-dimensional (3D) avatar form using Unity 3D.

Results: The VPs included four case scenarios that were congruent with the nursing undergraduates' semesters' learning objectives: (1) assessing the pain experienced by a pregnant woman, (2) taking the history of a depressed patient, (3) escalating a bleeding episode of a postoperative patient to a physician, and (4) showing empathy to a stressed-out fellow final-year nursing student. Challenges arose in terms of content development, technological limitations, and expectations management, which can be resolved by contingency planning, open communication, constant program updates, refinement, and training.

Conclusions: The creation of VPs to assist in nursing students' communication skills training may provide authentic learning environments that enhance students' perceived self-efficacy and confidence in effective communication skills. However, given the infancy stage of this project, further refinement and constant enhancements are needed to train the VPs to simulate real-life conversations before the official implementation.

(J Med Internet Res 2019;21(10):e14658) doi:10.2196/14658

KEYWORDS

artificial intelligence; communication; learning; nursing education; patients; technology; virtual reality

Introduction

Background

Effective communication skills are an integral part of the nursing profession and the foundation for high-quality nursing care [1]. Poor communication skills have been directly related to high turnover rates; low morale among nurses [2,3]; and poor patient outcomes such as medical errors, poor adherence to the treatment plans, and lower patient care satisfaction [4]. Effective communication between nurses and patients involving nurses' abilities to explain, listen, and empathize is necessary for successful outcomes in individualized patient care [5,6]. However, nursing students are often stressed over their lack of adequate skills to communicate effectively with patients and their family members [7,8]. This indicates a deficit in the availability of specialized communication training for nurses [7-9] and the ineffectiveness of the current communication skills training for nursing undergraduates through didactic lectures [10,11].

The need for communication skills training to be both participatory and experiential [11] led traditional nursing curricula to use simulated or standardized patients as tools to help students to develop clinical reasoning, patient communication, history taking, physical examination, and patient diagnosis skills [12,13]. Standardized patients are community members who are carefully recruited and trained to take on the characteristics of a real patient, and they provide students with opportunities of learning and assessments in simulated clinical environments [13]. However, the development and maintenance of a quality standardized patient program is costly and time-consuming [14]. Moreover, standardized patients are often subjected to feelings of anxiety, fatigue, physical discomfort, and biasness, which carry some reliability concerns [15,16]. Therefore, virtual patients (VPs) might be a more viable alternative.

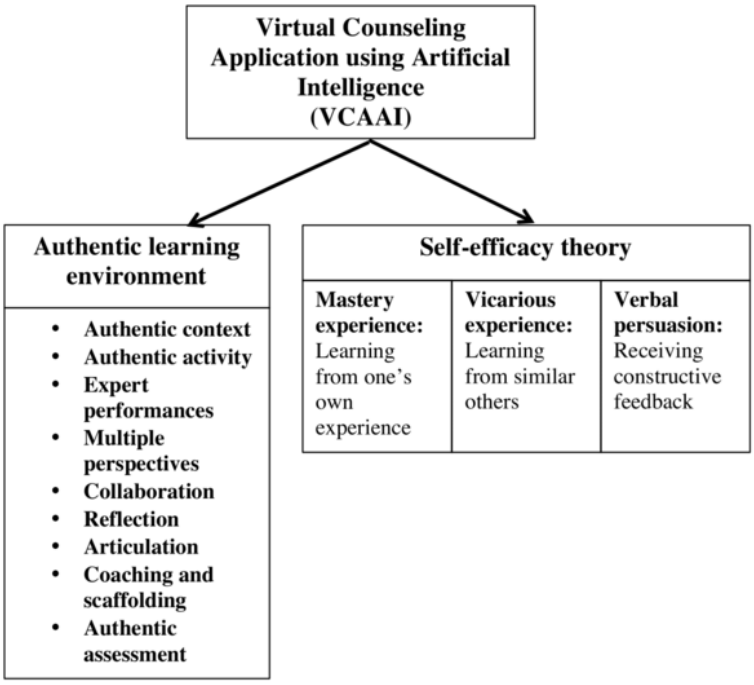
VPs are computer-based simulations of authentic clinical cases that allow users to interact with the system for the purpose of health care or medical trainings, education, or assessments [17,18]. Current uses of VPs in medical education are primarily to develop students' clinical reasonings, problem-solving skills, core or conceptual knowledge acquisitions, skills acquisitions, and affective characteristic developments (eg, professional competence) [19]. Reviews evaluating the effectiveness of VPs in medical education [17,20] have reported VPs as a cost-effective tool and as successful in facilitating clinical reasonings, communication skills, and ethical reasonings among students when used as an alternative or supplementary tool to existing curricula. However, in studies in which VPs have been used to teach and assess interview skills, students recognized the artificiality of such situations and did not demonstrate empathy or other important aspects of this skill [21,22]. Despite a few flaws in comparison with standardized patients, the use of VPs in medical education provided secured learning environments and opportunities for extensive repetitive practice with feedback and without negative consequences to real or standardized patients [21,23]. Apart from the acquisition of clinical knowledge and skills, VPs also provide students with opportunities for self-directed learning [24], which leads to reflection [25] and self-driven change [26,27]. Given their advantages and effectiveness for student learning outcomes, VPs are therefore more extensively used in medical education, although their use in nursing education is still limited [28].

Theoretical Framework

The theoretical frameworks that guided this study are Bandura's self-efficacy theory [29] and Herrington et al's [30] authentic learning concept. According to Bandura [29], self-efficacy is an individual's confidence about one's ability to carry out a behavior effectively, and this influences one's motivation and efforts in performing a specific task, individual goal setting, and perspective [31]. Therefore, self-efficacy is an essential component needed by nursing students to display effective communication skills in health care settings. Moreover, if students have higher perceived self-efficacy in their communication with patients, their family members, or other health care providers, they may be more inclined to initiate and engage in conversations, which can boost patient-provider and work relationships.

Bandura also mentioned the importance of mastery experience (learning from one's own experience) and verbal persuasion (receiving feedback on one's own performance) in enhancing an individual's self-efficacy [29]; therefore, these factors were taken into consideration when planning the VP case scenarios. In addition, 9 elements of authentic learning environments [30] consisting of authentic contexts, authentic activities, expert performances, multiple perspectives, collaboration, reflection, articulation, coaching and scaffolding, and authentic assessments will be introduced during the program planning to provide authentic exposure and encourage real-life learning to prepare nursing students for their future professional lives. The theoretical framework of this study is presented in Figure 1.

Figure 1. Theoretical framework for virtual counseling application using artificial intelligence.



Preliminary Study

This project is an extension of a preliminary study conducted in 2016 in which a blended learning approach was adapted to teach communication skills to year 1 nursing undergraduates [32,33]. The redesigned course retained face-to-face tutorials and replaced didactic face-to-face lectures with electronic lectures. Web-based quizzes, discussion forums, and reflection exercises were introduced to enhance students' engagement with the Web-based course material. To promote applications of the theoretical contents, face-to-face tutorials included role-play and problem-based learning using authentic clinical scenarios (ie, standardized patients) to simulate conversations between nurses, other health care professionals, patients, and their family members. Findings suggested that students who received the module via the blended learning approach had improved communication self-efficacy, better attitudes in learning communication skills, statistically significantly higher satisfaction, and better academic scores compared with participants from the previous cohort who had only didactic lectures [32]. Apart from enhanced learning, students also reported confidence boosts in handling similar situations in their year-end assessments, but they were less confident in retaining these learned skills and transferring it to real-life clinical settings during their end-of-semester clinical postings [33]. Stakeholders also mentioned deterioration in communication skills among year 3 and 4 nursing students, which indicates a necessity for additional resources to reinforce their communication skills. However, employing standardized patients, as such for year 1 students, to provide authentic communication training in subsequent years is expensive and resource extensive. Therefore, these findings motivated the development of VPs to fill in the previous gaps and provide nursing students with continuous authentic trainings in communication skills. Furthermore, the use of VPs also aligns with the university's goal of encouraging self-directed and autonomous learning among students using electronic platforms.

Methods

Overview

The Alice Lee Centre for Nursing Studies under the National University of Singapore (NUS) offers a 3-year (4 years for honors students), full-time Bachelor of Science (Nursing) program that is accredited by the Singapore Nursing Board. The course covers core modules, such as anatomy, physiology and physical assessment, pathophysiology, pharmacology and nursing practice, communication, and cultural diversity, and includes clinical practicums at tertiary hospitals that range from 2 weeks to 3 months. This project will be conducted with nursing undergraduates of the NUS who have completed the core module *Effective Communication for Health Professionals* (module code NUR1110) in the year 1 of their nursing courses. The 2-year study will follow these students in year 2 and year 3 consecutively by introducing VPs depicting real-life case scenarios at gradual difficulty levels before their end-of-semester clinical postings. A total of 4 VP case scenarios were developed for each semester on the following topics: (1) interviewing a pregnant woman with pain to solicit holistic history taking (year 2 semester 1); (2) history taking from a depressed patient (year 2 semester 2); (3) using a standardized approach such as situation, background, assessment, and recommendation (SBAR) to hand-off interdisciplinary communications (year 3 semester 1); and (4) showing empathy to a fellow nursing student (year 3 semester 2). Overall, the aim of this project was to develop and evaluate the use of VPs to better prepare nursing

undergraduates in communicating with real-life patients, family members, and other health care professionals during their clinical postings. The specific research questions we plan to answer in this project are as follows:

1. What is the effect of using VPs in enhancing nursing undergraduates' self-efficacy and attitudes toward learning communication skills?
2. Do the students who receive additional training using VPs perform better in their communication skills during their clinical postings compared with students who receive standard training?
3. What are the levels of outcomes of students' self-efficacy and attitudes toward learning communication skills at pretest (year 2 semester 1 before receiving VP training), posttest 1 (last day of clinical posting year 2 semester 1), posttest 2 (last day of clinical posting year 2 semester 2), posttest 3 (last day of clinical posting year 3 semester 1), and posttest 4 (last day of clinical posting year 3 semester 2)?
4. What are the changes in self-efficacy and attitudes toward communication skills scores over time (pretest and posttests 1 to 4)?
5. What are the students' experiences in receiving additional training using VPs before their clinical postings?

The aim of this paper was, therefore, to provide a detailed breakdown on the development process of the virtual counseling application using artificial intelligence (VCAAI) for communication skills training in nursing education and to highlight recommended resolutions and challenges faced to inform future research.

Design and Development

The research team involved in the design and development of the VCAAI comprises clinical nurses, nurse educators, and information technology (IT) experts. On the basis of the vast experience of the research team, the detailed dialogue flows with mind maps were written at the initial stage. The Master Interview Rating Scale (MIRS) provided a framework to develop the scenarios. The MIRS was designed to teach effective communication between health care practitioners and patients. It has been used in medical education over the past decades [34]. On the basis of the 15 basic items in the MIRS framework, the team created a coding schema to classify the characteristics of a patient's interview questions into 1 of the following categories: open-ended, close-ended, empathetic statements, information gathering, and the patient's perspective.

The nursing research team then worked closely with the technology team from the NUS Information Technology department to develop a voice chatbot learning system. The voice chatbot was powered by an artificial intelligence (AI) using the natural language processing engine Google Cloud's Dialogflow. The AI uses artificial neural networks just like human intelligence to learn from varying situations to recognize, classify, and predict responses based on analysis by machines such as computer systems [35]. In this study, a limited memory AI [36] was used to mimic the human characteristics of standardized patients used in simulation training for communication skills. The standardized patient's conversations were further visualized in a three-dimensional (3D) avatar form, characterized by natural, nonverbal gestures to elicit more engaging connections as well as to have more life-like realism. It was later further integrated and visualized into a 3D avatar form to mimic human conversations (both verbal and nonverbal) by leveraging Unity 3D, which is a popular 3D development platform. According to Hintze et al [36], a limited memory AI uses past experiences to inform future decisions. As such, our 3D avatar (VP) was trained using Google Cloud's Dialogflow processing engine to store memories of potential conversations. Supervised learning by the machines (VPs) was modeled directly based on the research team's observations and an analysis of live standardized patient interactions (video recordings) with the nursing undergraduates from our previous preliminary study.

Case Scenarios

Students can interact with the VPs through 4 case scenarios, which were developed by members of the clinical nursing team with input from nursing students with a clinical background. These scenarios are based on authentic clinical cases (adapted from real-life clinical case studies), focusing mainly on communication aspects. The topic details and student objectives for the 4 scenarios described in Textbox 1.

The four scenarios were developed based on the growing clinical needs of nursing undergraduates. Each scenario is matched with the core modules students take during their undergraduate years. For example, the *pregnant woman* scenario is matched with the core module *Women and Child Health* that students take in year 2 semester 1. The *depressed patient* scenario is matched with the *Mental Health* module that students take in year 2 semester 2. The *SBAR* scenario is matched with the year 3 semester 1 module *Operating Theatre Nursing*, and the last *empathy* scenario is matched with the final year 3 semester 2's clinical postings before students transit to clinical settings and registration with the Singapore Nursing Board.

Textbox 1. Case scenarios.

Scenario 1: A 32-year-old pregnant woman in her third trimester experienced pain along the side of her belly while heading to the market in the morning. She was later diagnosed with Braxton Hicks because of a lack of consistent contractions and no increase in contraction strength. In addition, she is prone to menstrual cramps and has a history of miscarriage, which has rendered her very anxious, scared, and worried about her pregnancy. She does not have any children currently, and her spouse is often very busy.

- Objective: Students are required to interview the patient and find out more about her pain. Using the pneumonic *c*haracteristics of the pain, *o*nset of the pain, *l*ocation of the pain, *d*uration of the pain, severity of the pain, *p*recipitating factors that make the pain worse or make her feel better, and *a*ssociated symptoms (COLDSPA), students will communicate with the pregnant woman to obtain a holistic idea about her pain and mutually develop her care plan to allay her anxiety.

Scenario 2: A 34-year-old, single, Chinese male lorry driver self-admitted himself to the hospital because of a relapse of depressive symptoms. He is in a depressed mood and has a closed-off demeanor, yet is fidgety and actively avoids eye contact. He has a history of major depressive disorder and was previously admitted because of a failed suicide attempt. He has undergone cognitive behavioral therapy and has been prescribed fluoxetine. Although his condition was stable at discharge, he is noncompliant with his prescribed regimen of fluoxetine and now complains of a loss of appetite, underproductivity, avolition, and reclusiveness.

- Objective: Students are required to establish a working rapport with the patient and elicit background information from a biological, psychological, and sociological perspective using the MIRS and relevant components of COLDSPA. Students will then proceed to develop a care plan based on the information they gather.

Scenario 3: A 42-year-old male patient was admitted to the general surgical ward after an operation (post appendectomy) 3 days ago. The student is assigned to change the patient's operative wound dressing, but the dressing is soaked in blood, and he appears to be slightly pale, lethargic, and distressed.

- Objective: Students need to use the pneumonic SBAR to update the patient's condition to the physician as well as to demonstrate clinical reasoning and appropriately prioritize areas of concern in patient care.

Scenario 4: The student user will take the role as a final-year nursing student communicating with another student (VP) during her final preregistration clinical posting period. The student will be very stressed because of escalating demands and clinical workloads.

- Objective: The student user will speak to this stressed student to find out more about her stressors and help her to reflect on how she can cope better. This scenario focuses on the principles of *Showing Empathy* using the pneumonic *N*aming emotion, express *U*nderstanding, showing *R*ecognition, and offering *S*upport (NURS) and on helping someone obtain his or her own perspective using the pneumonic *I*deas or beliefs of cause of the situation, how the situation is affecting daily *F*unction, and what *E*xpectations one has from the interviewer over the situation (IFE). These pneumonics have been taught in the *Effective Communication Among Health Professionals* core module, and the students should be well-versed in their use.

User Interface

The VCAAI user interface was designed to be straightforward and user friendly with realistic features and detailed instructions to guide the students and optimize their user experiences. At the start of every session, students will arrive at the case scenario selection menu to choose 1 of the 4 scenarios mentioned previously (Figure 2).

The cases of the pregnant woman and the depressed patient were selected for illustrative purposes. Students will be first provided with introductory instructions to begin conversing with the completed 3D voice chatbot (Figure 3), after which they will be given interview objectives and instructions depicting their roles and relationships to the patient, which are guided closely by the context of the case (Figure 4).

Before interaction with the VP, students must complete a speech recognition training involving the main keywords related to the selected case scenario (Figure 5).

Upon successful pronunciations of a few keywords, students can proceed to interview the VP as per the nursing conversation guidelines (MIRS, COLDSPA, SBAR, NURS, and IFE). They will be presented with a first-person perspective of the patient avatar in 3D with his or her demographic profile and a list of potential questions to interact with the VP (Figure 6).

Figure 2. Case scenario selection menu for the virtual counseling application using artificial intelligence.

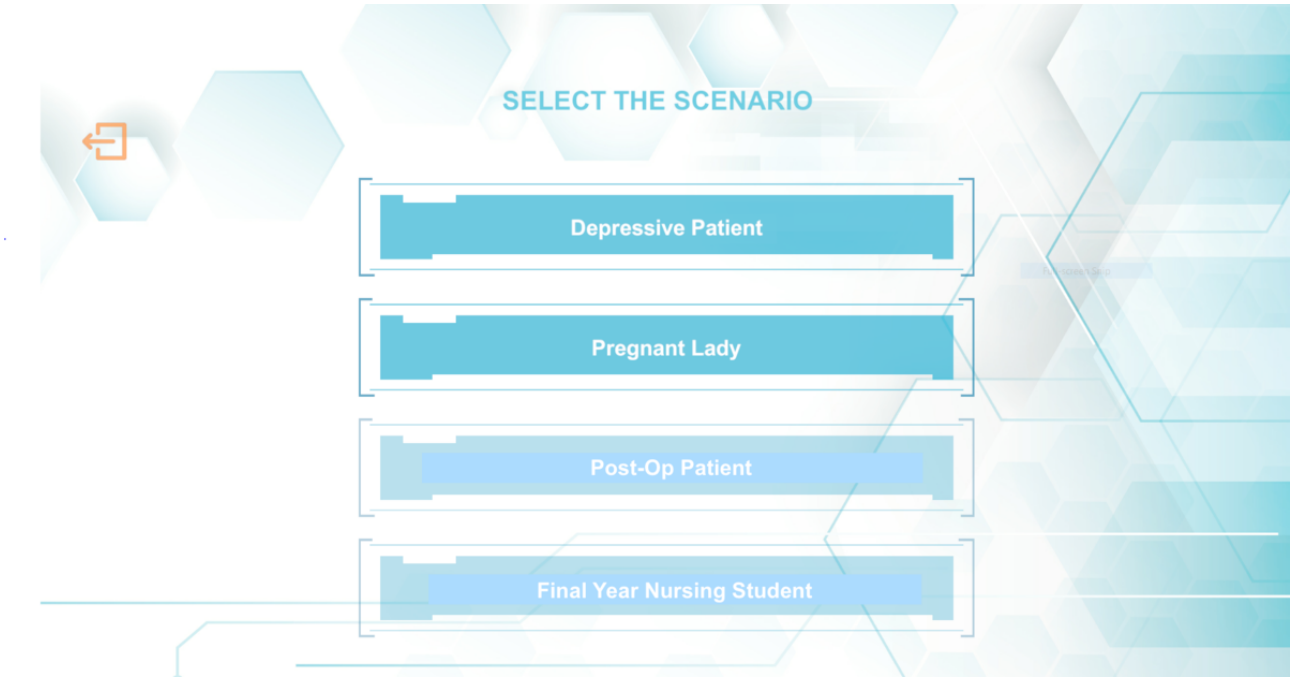


Figure 3. Technical voice input instructions on how to initiate a conversation with the virtual patient.

Figure 4. Assessment instructions and objectives for students.

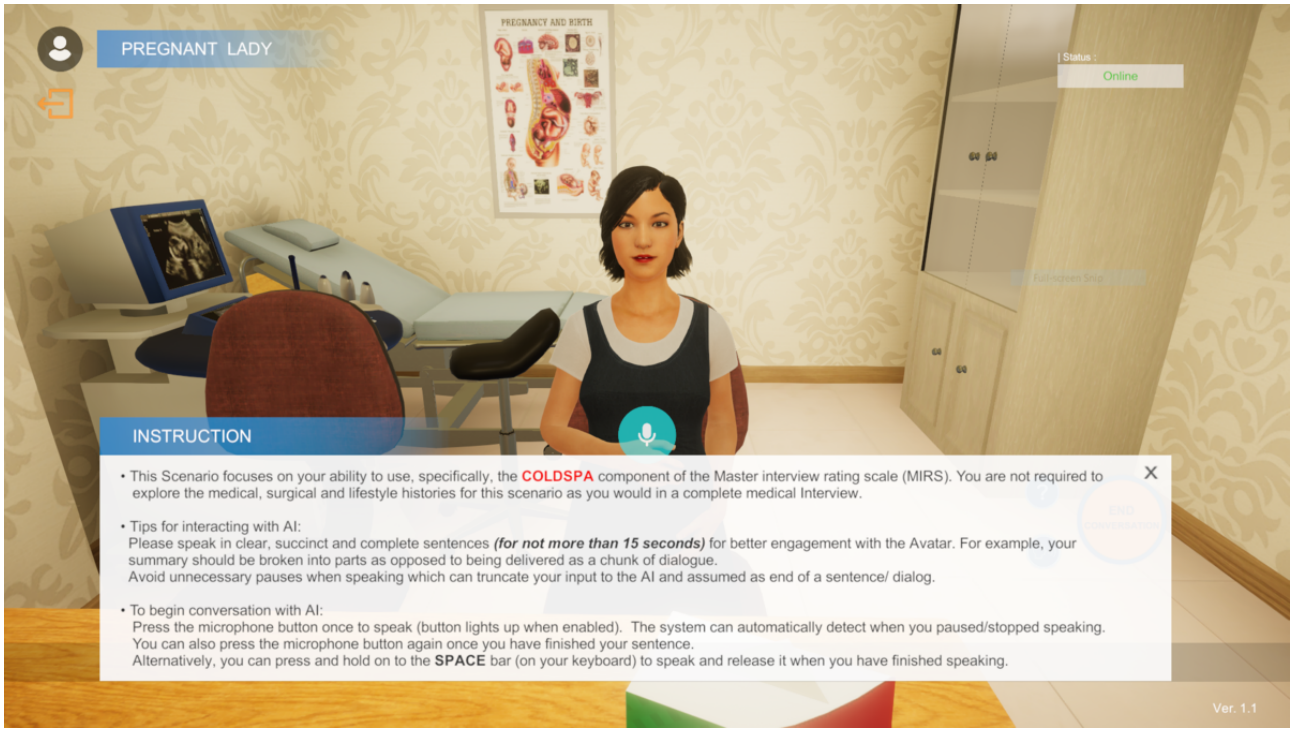


Figure 5. List of case scenario–specific keywords provided for voice recognition training.

Figure 6. Case scenario of a depressed patient.

The list of questions is compiled from reviewing previously video-recorded interactions between standardized patients and nursing undergraduates while they interview these patients. These interviews are part of the year 1 nursing undergraduate core module titled *Effective Communication Among Health care Professionals' Final Assessment*. The users have complete freedom to determine the sequence of the interviews. Nonverbal cues, such as the head-in-the-hands gesture, and a realistic clinical environment with relevant posters and equipment are incorporated to enhance the simulation of a real-life scenario. The scenario will end with a case summary discussing the main points shared, a performance checklist that serves as a formative feedback on their communication skills and interview processes, and a closure to recommend the next series of actions to take (Figure 7).

Figure 7. Performance checklist at the end of the assessment.

Results

To test and validate the content, the VCAAI will be implemented on approximately 150 year 2 and year 3 nursing undergraduates. A user acceptance test (UAT) will be performed to evaluate the program's effectiveness on nursing students' performances, attitudes, and perceived self-efficacy pertaining to communication skills. Only nursing undergraduates who have completed the redesigned NUR1110 (Effective Communication for Health Professionals) core module will receive additional training via VPs in each semester (2 semesters per year) of year 2 and year 3 before they go for their clinical postings. Students will have unlimited access to the VPs (available scenarios depend on which semester the student is in) by logging in through the school portal. At the end of the semester, students' performances, attitudes, and self-perceived self-efficacy in terms of communication skills will be measured using self-administered questionnaires, and the consolidation of students' experiences and feedback will be done through focus group discussions. Approximately 20 clinical instructors (the exact number will be based on data saturation) supervising the nursing students during their end-of-semester clinical postings will also be interviewed for their perspectives of the students' communication skills abilities. Subsequently, VPs users' comments and ambiguities generated from the focus group discussions will be incorporated into the VCAAI program to further refine and optimize the user experience. Only then can the VPs be officially implemented and made available to all nursing students on the academic portal.

Discussion

Throughout the course of development to the execution of the scenarios of the VCAAI program, the team met with various technological limitations and challenges to prove the effectiveness of the VPs, to imitate a standardized patient, and to provide smooth role-playing experiences for students. The context of these challenges and resolutions are categorized as follows: (1) content development, (2) technological limitations, (3) expectation management and contingency planning.

Content Development

Content development entails permutation development, presentation of permutation, the relay of permutation to program designers for development, and the ability to predict possible dialogue permutations. Each of these aspects is important to ensure that the user experience with the VPs is fulfilling and beneficial.

A prominent issue that arose at the initial stage was that programmers and content developers spoke in different *languages* because of a lack of expertise in each other's field. Content developers were focused on nursing communication skills, but programmers were unfamiliar with nursing concepts, conversation guidelines (MIRS, COLDSPA, SBAR, NURS, and IFE), and what a conversation between a nurse and a patient would be like. In addition, content developers did not know what information the programmers required and were unfamiliar with the technical terms used by programmers, and hence, they had difficulty translating their knowledge into *programmers' language*. This difficulty in communication put restraints on the instructions and expectations of the VPs. Both teams should be more mindful when using technical jargon so that everyone can be on the same page, and technicalities should clearly be spelled out so that misunderstandings can be prevented. For example, the content developers can provide an information sheet that clearly explains to the IT team and vendor the common terms and expectations of health care communications in various scenarios. Visual diagrams and specific examples (eg, mind maps, drawings, and sample dialogues) were also useful in helping the programmers understand the content developers' expectations of conversations. Such information sheets should be provided by both teams, which will give each other greater insights of each field, and this document should be kept accessible for easy reference when working individually. In addition, for the content developers to provide adequate information for the program developers, the programmers must create specific templates and constantly scrutinize the data before the data can be translated into the VPs. Moreover, administrative details such as the labeling of the files and the color coding of the data had to be mutually agreeable and adhered to by both parties.

To bridge the gap of expertise and information between content designers and IT experts, it will help to have a team member who encompasses expertise in both content design and IT. Frequent check-backs of the program process is needed through weekly face-to-face meetings. The communication styles of the nursing team (content developers) and the IT experts differ, and the keywords used to describe similar concepts vary, leaving room for misunderstandings and confusion, which can eventually cause distrust and frustration toward other team members. To resolve this issue, content developers and programmers maintained open communication and corresponded closely through mutual platforms (ie, emails and WhatsApp). Weekly face-to-face meetings were also conducted to prevent miscommunication through emails and WhatsApp messages. However, with the development of subsequent scenarios, there were marked improvements in terms of communication between team members and ease in the development process, which could be attributed to familiarity with the project by both teams.

To further improve the workflow between content developers and IT experts, a direct discussion with the actual program developers may be more effective than discussion with a representative. In this project, with special circumstances, the program developers were in a different country with a different time zone from the content developers, and the weekly meetings were communicated via a representative of the IT expert team who used an online chatting platform to relay the team's message to the developers. This left room for miscommunication. Hence, it would be ideal for the program developers to be present more frequently during meetings and test runs so that they can experience the realities of the way the team communicates as well as bot performances with a user. It will also paint a clearer picture for them when they make amendments. However, given the tight budget, it was not feasible to fly them in on a weekly or even biweekly basis. The time zone differences also made organizing Skype calls difficult. For future collaborations, the availability and presence of

on-the-ground developers can be considered before engaging an IT expert team.

Another challenge faced by the team was their inability to predict the possible intent of the conversations. Although content developers were focused on the components of nursing conversation guidelines (MIRS, COLDSPA, SBAR, NURS, and IFE) and various possible imputations, they failed to consider other conversational intents that went beyond the guidelines, such as small talk and giving advice. As students go beyond the scope of the guidelines, the content developers have major difficulties keeping up with the infinite possibility of inputs. This renders the conversations with the VPs disruptive and thus fails to effectively provide an immersive experience of role-playing for students. A resolution of this issue involved including senior nursing students, who had previously taken this communication module, in the VP testing to provide content developers with more permutations to broaden the VPs' abilities to have natural conversations. However, the team was still met with the issue of keeping up with unexpected inputs by students, as seen in the UATs. Therefore, to encode the VPs to reflect the complexity and multifaceted nature of a person, programmers will have to constantly update the VPs using past inputs provided by the users. This process needs to be ongoing for at least one school semester, given the complexity of a natural conversation. Despite all the enhancements, there will always be some limitations for the VPs to account for because of countless unpredictable permutations. A continuous training of the VP will be needed to exhaust the maximum amount of data permutations with as many UATs as possible to report visible improvements only after each iteration of such a training in a longitudinal study. As such, VPs may be more effective with structured conversations and predictable inputs. For example, in the third scenario, as SBAR is more structural than conversational compared with the rest of the scenarios, the content design was relatively more linear and straightforward. Furthermore, playing as a fellow health care colleague receiving information from the user, the bot's responses were often acknowledgment statements such as *okay* and *alright*, making content design a lot simpler. Upon testing within the team members, the conversation flow was indeed a lot more predictable and compartmentalized, and the majority of the responses from the program were appropriate.

Technological Limitations

The VP is unable to adapt to the context of the conversation as the basis of communication for the VP is mostly reactive to the input, and it retrieves only a specific response tagged to the permutation. For example, in the first scenario:

> *Student: On a scale of zero to ten, can you rate your pain [belly pain]?*
>
> *VP: Seven out of ten.*
>
> *Student: Do you have any other pain?*
>
> *VP: My back aches a little too.*
>
> *Student: On a scale of zero to ten, can you rate your pain [back pain]?*
>
> *VP: Seven out of ten [although the intended pain score for back pain was two out of ten to trigger prioritizing issues].*

As seen, the keyword *pain score* was tagged with the answer *seven out of ten* as the belly pain was the main chief complaint. The VP was unable to adapt to the context, and this resulted in difficulties for content developers in tagging keywords to responses as similar questions can refer to different issues that require different responses.

Another example is the fourth scenario, which explores the use of IFE and NURS. As the program recognizes keywords to determine the appropriate responses, users may say *you seem stressed* with the intention of naming the emotion, and the keywords here are *seem* and *stressed*. Users may also say *don't be stressed* with the intention of consoling, and the keywords are *don't be* and *stressed*. With very similar keywords, the program may sometimes respond inappropriately, switching up responses for these 2 variations. With limited knowledge on how the program recognizes keywords or associates keywords together, content creators sometimes choose keywords inappropriately, resulting in the program being unable to fulfill its function. Such loopholes are easily missed out unless variations of the keywords are tested out with the program, the transcripts of the test runs are reviewed, and the content is amended afterward. This is manpower-intensive and time-consuming and may not be the most effective way to move forward. The possibility of the program learning variations and logic can be explored through predictions of speech and language.

Another key concern was that not all computers (eg, Mac, older versions of Windows, and Windows without Cortana or a graphic card) are compatible with the application, and certain earpieces or microphones do not work with the application. This led to a poor recognition of voice inputs and testers' dictions, especially when a diction is heavily influenced by culture and content and the software used to develop the VPs uses American English. The software's inability to detect, recognize, and translate speech to text effectively poses a challenge. There are various possible factors such as background noises and poor microphone qualities, but through the team's observation, the students' dictions were the primary issue. As Cortana (Microsoft Corporation), a voice recognition software, uses either American English or British English, if the students are more comfortable with Singlish (English-based creole or patois spoken colloquially in Singapore), the program will have difficulty recognizing their pronunciations or speech patterns, resulting in translation failures. For example, in addition to unique and complex medical terms (eg, laparoscopic appendectomy and postoperation day), when students attempt to mimic the Singapore setting (ie, by naming various patients with common Singaporean names), the translations can be inaccurate. Moving forward, to counter text translation difficulties, audio-recordings of real Singaporean users will be incorporated into the program to enhance text translation abilities and to eliminate some errors from varying accents and pronunciation differences between Singaporean English and American English. In addition, we have invested in good quality microphones for the students during the UATs and have consulted with software developers to use the latest

voice-training feature. Early setups and test runs of the computers and VPs by the team are crucial before test runs with students.

This program of virtual learning aims to enhance learning and mimic reality to the technology's best capacity. Although environmental and nonverbal cues were taken into consideration when programming the VPs' settings and behaviors, these aspects of communication can be further refined during interactions between students and VPs by including action-related selections for users to select or active reflection elements that allow the users to reflect on time points in the conversation continuums, during which nonverbal communication and environmental factors could have made a difference. For instance, in the third scenario, although SBAR can be structural, much thought had to be put into the scenarios so that the conversations do not seem robotic or linear. The program must carry a certain caliber of conversational capacities and physical attractiveness to enchant users. Conversational elements are often unpredictable and arise from real user test runs of the program. Like previous scenarios, with a small team being familiar with the nursing context and conversational skills, variations may be limited. Future studies that are set in diverse and multicultural environments should consider including other languages or dialect inputs for their VP programs to simulate a more realistic hospital case setting for students.

Expectations Management and Contingency Planning

A possible setback to the teamwork is the misunderstanding of expectations toward other expertise. The content creators were unsure of the processes needed to incorporate changes and the logic of the program operating system, whereas the IT team were unsure of the expectations of the outcome of the program. As the program is still in its preliminary stage, expectation management of the team is essential to keep the team grounded and prepared for potential hiccups. Establishing open and clear communication, setting realistic team and user expectations and deadlines, and having an administrative figure to keep track of deadlines and to keep everyone on track will help to facilitate an effective operation of the task and to ensure that all members are equally involved in the end goal. Meeting minutes should also be taken at the end of every meeting and sent to all members for reference. A consolidation of the deliverables should also be readily available and should be revisited before every subsequent meeting. It is important that before every UAT, clear instructions are given to the students that the purpose of using the VPs is to provide additional training to practice communication skills and not to replace actual or standardized patients. This may provide them with a sense of realism about the actual intent of the VCAAI program, which will thus avoid unnecessary comparisons and follow-on frustrations.

During the development stage, the team should prepare a blueprint by breaking the project into smaller parts, such as content development and IT development. This includes discussing possible exhaustive permutations during content development and having weekly meetings between the IT developers and the content developers to clarify any doubts and to document the process and progress of the project until its completion. There should be exigency plans throughout the development and execution phases of the VPs and UATs. Moreover, 2 of the most important lessons that were learned during the development phase of the VPs were to factor enough time for the content development and training of the VPs and to be prepared with alternatives. We also foresee challenges of recruiting and engaging students during the UATs as the test runs will take place over 4 different semesters of the nursing undergraduate program. Having strategic plans such as sending multiple reminders to the students before the UATs, providing them options of engaging in UATs even from home by downloading the provided links, and planning the UATs' dates to match their academic timetables to avoid clashes with their other classes will be helpful.

Conclusions

The adoption and creation of VP simulations to hone nursing students' communication and interview skills with patients, health care providers, and colleagues will not only provide truly unique and authentic learning experiences for students but also potentially enhance their perceived self-efficacy and confidence in effective communication skills. However, given the infancy stage of this project, further refinements and constant enhancements are needed to train the VPs, such as by increasing possible permutations and improving speech-to-text translations to facilitate more realistic conversations. Future developers and users of the VCAAI can learn from the reflections shared by the authors to avoid similar pitfalls and enhance the user experience.

Acknowledgments

The authors would like to thank the Ministry of Education, Tertiary Research Funds (MOE2017-TRF-008) for funding this project. The authors also extend their thanks to the NUS Information Technology department for managing the development of the system. Finally, the authors extend special thanks to the NUS's nursing students for their invaluable support to this project.

Conflicts of Interest

None declared.

References

1. Mccabe C, Timmins F. Communication Skills For Nursing Practice. Second Edition. Basingstoke, United Kingdom: Palgrave Macmillan; 2013.

2. Brinkert R. A literature review of conflict communication causes, costs, benefits and interventions in nursing. J Nurs Manag 2010 Mar;18(2):145-156. [doi: 10.1111/j.1365-2834.2010.01061.x] [Medline: 20465742]
3. Vessey JA, Demarco R, DiFazio R. Bullying, harassment, and horizontal violence in the nursing workforce: the state of the science. Annu Rev Nurs Res 2010;28:133-157. [doi: 10.1891/0739-6686.28.133] [Medline: 21639026]
4. Foronda C, MacWilliams B, McArthur E. Interprofessional communication in healthcare: an integrative review. Nurse Educ Pract 2016 Jul;19:36-40. [doi: 10.1016/j.nepr.2016.04.005] [Medline: 27428690]
5. Kourkouta L, Papathanasiou IV. Communication in nursing practice. Mater Sociomed 2014 Feb;26(1):65-67 [FREE Full text] [doi: 10.5455/msm.2014.26.65-67] [Medline: 24757408]
6. Duffy FD, Gordon GH, Whelan G, Cole-Kelly K, Frankel R, Buffone N, Participants in the American Academy on Physician and Patient's Conference on Education and Evaluation of Competence in Communication and Interpersonal Skills. Assessing competence in communication and interpersonal skills: the Kalamazoo II report. Acad Med 2004 Jun;79(6):495-507. [doi: 10.1097/00001888-200406000-00002] [Medline: 15165967]
7. Krimshtein NS, Luhrs CA, Puntillo KA, Cortez TB, Livote EE, Penrod JD, et al. Training nurses for interdisciplinary communication with families in the intensive care unit: an intervention. J Palliat Med 2011 Dec;14(12):1325-1332 [FREE Full text] [doi: 10.1089/jpm.2011.0225] [Medline: 22132740]
8. Villagran M, Goldsmith J, Wittenberg-Lyles E, Baldwin P. Creating comfort: a communication-based model for breaking bad news. Commun Edu 2010 Jul;59(3):220-234. [doi: 10.1080/03634521003624031]
9. Wittenberg-Lyles EM, Goldsmith J, Sanchez-Reilly S, Ragan SL. Communicating a terminal prognosis in a palliative care setting: deficiencies in current communication training protocols. Soc Sci Med 2008 Jun;66(11):2356-2365. [doi: 10.1016/j.socscimed.2008.01.042] [Medline: 18321625]
10. Chant S, Tim, Randle J, Russell G, Webb C. Communication skills training in healthcare: a review of the literature. Nurse Educ Today 2002 Apr;22(3):189-202. [doi: 10.1054/nedt.2001.0690] [Medline: 12027600]
11. Parry R. Are interventions to enhance communication performance in allied health professionals effective, and how should they be delivered? Direct and indirect evidence. Patient Educ Couns 2008 Nov;73(2):186-195. [doi: 10.1016/j.pec.2008.05.029] [Medline: 18768287]
12. Brender E, Burke A, Glass RM. JAMA patient page. Standardized patients. J Am Med Assoc 2005 Sep 7;294(9):1172. [doi: 10.1001/jama.294.9.1172] [Medline: 16145035]
13. Churchouse C, McCafferty C. Standardized patients versus simulated patients: is there a difference? Clin Simulat Nurs 2012 Oct;8(8):e363-e365. [doi: 10.1016/j.ecns.2011.04.008]
14. Bosse HM, Nickel M, Huwendiek S, Schultz JH, Nikendei C. Cost-effectiveness of peer role play and standardized patients in undergraduate communication training. BMC Med Educ 2015 Oct 24;15:183 [FREE Full text] [doi: 10.1186/s12909-015-0468-1] [Medline: 26498479]
15. Plaksin J, Nicholson J, Kundrod S, Zabar S, Kalet A, Altshuler L. The benefits and risks of being a standardized patient: a narrative review of the literature. Patient 2016 Feb;9(1):15-25. [doi: 10.1007/s40271-015-0127-y] [Medline: 26002043]
16. Kroboth FJ, Hanusa BH, Parker S, Coulehan JL, Kapoor WN, Brown FH, et al. The inter-rater reliability and internal consistency of a clinical evaluation exercise. J Gen Intern Med 1992;7(2):174-179. [doi: 10.1007/bf02598008] [Medline: 1487766]
17. Consorti F, Mancuso R, Nocioni M, Piccolo A. Efficacy of virtual patients in medical education: a meta-analysis of randomized studies. Comput Educ 2012 Nov;59(3):1001-1008. [doi: 10.1016/j.compedu.2012.04.017]
18. Ellaway R, Poulton T, Fors U, McGee JB, Albright S. Building a virtual patient commons. Med Teach 2008;30(2):170-174. [doi: 10.1080/01421590701874074] [Medline: 18464142]
19. Midik O, Kartal M. The use of virtual patients in medical education. Marmara Med J 2015;28(2):63. [doi: 10.5472/MMJoa.2802.02]
20. Cook DA, Triola MM. Virtual patients: a critical literature review and proposed next steps. Med Educ 2009 Apr;43(4):303-311. [doi: 10.1111/j.1365-2923.2008.03286.x] [Medline: 19335571]
21. Deladisma AM, Cohen M, Stevens A, Wagner P, Lok B, Bernard T, Association for Surgical Education. Do medical students respond empathetically to a virtual patient? Am J Surg 2007 Jun;193(6):756-760. [doi: 10.1016/j.amjsurg.2007.01.021] [Medline: 17512291]
22. Raij AB, Johnsen K, Dickerson RF, Lok BC, Cohen MS, Duerson M, et al. Comparing interpersonal interactions with a virtual human to those with a real human. IEEE Trans Vis Comput Graph 2007;13(3):443-457. [doi: 10.1109/TVCG.2007.1036] [Medline: 17356212]
23. Stevens A, Hernandez J, Johnsen K, Dickerson R, Raij A, Harrison C, et al. The use of virtual patients to teach medical students history taking and communication skills. Am J Surg 2006 Jun;191(6):806-811. [doi: 10.1016/j.amjsurg.2006.03.002] [Medline: 16720154]
24. Cendan J, Lok B. The use of virtual patients in medical school curricula. Adv Physiol Educ 2012 Mar;36(1):48-53 [FREE Full text] [doi: 10.1152/advan.00054.2011] [Medline: 22383412]
25. Boyd EM, Fales AW. Reflective learning: key to learning from experience. J Humanist Psychol 1983 Apr 1;23(2):99-117. [doi: 10.1177/0022167883232011]

26. Ridley DS, Schutz PA, Glanz RS, Weinstein CE. Self-regulated learning: the interactive influence of metacognitive awareness and goal-setting. J Experiment Educ 1992;60(4):293-306. [doi: 10.1080/00220973.1992.9943867]
27. Zimmerman BJ. Self-regulated learning and academic achievement: an overview. Educ Psychol 1990 Jan;25(1):3-17. [doi: 10.1207/s15326985ep2501_2]
28. Ellaway RH, Davies D. Design for learning: deconstructing virtual patient activities. Med Teach 2011;33(4):303-310. [doi: 10.3109/0142159X.2011.550969] [Medline: 21456988]
29. Bandura A. Self-Efficacy: The Exercise Of Control. New York, United States: Worth Publishers; 1997.
30. Herrington J, Reeves TC, Oliver R. A Guide to Authentic e-Learning. Oxfordshire: Routledge; 2009.
31. Wood R, Bandura A. Impact of conceptions of ability on self-regulatory mechanisms and complex decision making. J Pers Soc Psychol 1989 Mar;56(3):407-415. [doi: 10.1037//0022-3514.56.3.407] [Medline: 2926637]
32. Shorey S, Kowitlawakul Y, Devi MK, Chen HC, Soong SK, Ang E. Blended learning pedagogy designed for communication module among undergraduate nursing students: a quasi-experimental study. Nurse Educ Today 2018 Feb;61:120-126. [doi: 10.1016/j.nedt.2017.11.011] [Medline: 29197264]
33. Shorey S, Siew AL, Ang E. Experiences of nursing undergraduates on a redesigned blended communication module: a descriptive qualitative study. Nurse Educ Today 2018 Feb;61:77-82. [doi: 10.1016/j.nedt.2017.11.012] [Medline: 29190535]
34. Thomas TA. StudyLib. 2010. MIRS: Master Interview Rating Scale The Clinical Conversation and Interview: The Training Manual for Standardized Patients URL: https://studylib.net/doc/7245886/the-medical-conversation-and-interview [accessed 2019-09-10]
35. Miller T. Explanation in artificial intelligence: insights from the social sciences. Artif Intell 2019;267:1-38. [doi: 10.1016/j.artint.2018.07.007]
36. Hintze A, Kirkpatrick D, Adami C. The Structure of Evolved Representations Across Different Substrates for Artificial Intelligence. In: Proceedings of the Conference on Artificial Life. 2018 Presented at: ALIFE'18; July 23-27, 2018; Tokyo, Japan p. 388-395. [doi: 10.1162/isal_a_00076]

Abbreviations

3D: three-dimensional
AI: artificial intelligence
COLDSPA: Characteristics of the pain, onset of the pain, location of the pain, duration of the pain, severity of the pain, precipitating factors that make the pain worse or make her feel better, and associated symptoms
IFE: Ideas or beliefs of cause of the situation, how the situation is affecting daily Function, and what Expectations one has from the interviewer over the situation
IT: information technology
MIRS: Master Interview Rating Scale
NURS: Naming emotion, express Understanding, showing Recognition, and offering Support
NUS: National University of Singapore
SBAR: situation, background, assessment, and recommendation
UAT: user acceptance test
VCAAI: virtual counseling application using artificial intelligence
VP: virtual patients

Edited by N Zary; submitted 08.05.19; peer-reviewed by J Sharit, E Frontoni; comments to author 10.07.19; revised version received 02.08.19; accepted 31.08.19; published 29.10.19.

Please cite as:
Shorey S, Ang E, Yap J, Ng ED, Lau ST, Chui CK
A Virtual Counseling Application Using Artificial Intelligence for Communication Skills Training in Nursing Education: Development Study
J Med Internet Res 2019;21(10):e14658
URL: https://www.jmir.org/2019/10/e14658
doi:10.2196/14658
PMID:31663857

Original Paper

The Psychometric Properties of the Chinese eHealth Literacy Scale (C-eHEALS) in a Chinese Rural Population: Cross-Sectional Validation Study

Zhihao Ma[1], PhD; Mei Wu[2], PhD

[1]Computational Communication Collaboratory, School of Journalism and Communication, Nanjing University, Nanjing, China

[2]Department of Communication, Faculty of Social Sciences, University of Macau, Macau, China

Corresponding Author:
Zhihao Ma, PhD
Computational Communication Collaboratory
School of Journalism and Communication
Nanjing University
Room 362, Zijin Building, Nanjing University (Xianlin Campus), 163 Xianlin Road, Qixia District, Jiangsu
Nanjing, 210023
China
Phone: 86 17561538460
Email: redclass@163.com

Abstract

Background: The eHealth Literacy Scale (eHEALS) is the most widely used instrument in health studies to measure individual's electronic health literacy. Nonetheless, despite the rapid development of the online medical industry and increased rural-urban disparities in China, very few studies have examined the characteristics of the eHEALS among Chinese rural people by using modern psychometric methods. This study evaluated the psychometric properties of eHEALS in a Chinese rural population by using both the classical test theory and item response theory methods.

Objective: This study aimed to develop a simplified Chinese version of the eHEALS (C-eHEALS) and evaluate its psychometric properties in a rural population.

Methods: A cross-sectional survey was conducted with 543 rural internet users in West China. The internal reliability was assessed using the Cronbach alpha coefficient. A one-factor structure of the C-eHEALS was obtained via principal component analysis, and fit indices for this structure were calculated using confirmatory factory analysis. Subsequently, the item discrimination, difficulty, and test information were estimated via the graded response model. Additionally, the criterion validity was confirmed through hypothesis testing.

Results: The C-eHEALS has good reliability. Both principal component analysis and confirmatory factory analysis showed that the scale has a one-factor structure. The graded response model revealed that all items of the C-eHEALS have response options that allow for differentiation between latent trait levels and the capture of substantial information regarding participants' ability.

Conclusions: The findings indicate the high reliability and validity of the C-eHEALS and thus recommend its use for measuring eHealth literacy among the Chinese rural population.

(J Med Internet Res 2019;21(10):e15720) doi:10.2196/15720

KEYWORDS

eHealth literacy; eHEALS; psychometrics; classical test theory; item response theory

Introduction

China has the world's largest population of internet users, who also frequently access medical resources over the internet [1]. In fact, 26.6% of Chinese internet users use the internet to access online medical service [1]. Internet users adopt online medical platforms and related applications for evaluation of doctors [2], medical inquires [3], and health management [4]. The Chinese government and health practitioners have recognized challenges and opportunities, given the increasing adoption and diffusion of information communication technologies (ICTs) in the health

care system [5]. For example, limited health literacy has been shown to pose problems for Chinese residents who access online health resources [6] and has even led to avoidable medical tragedies [7]. Promoting public health literacy in today's information age is an urgent need in China. Therefore, in 2019, the National Health Commission of the People's Republic of China proposed the *Health China Campaign (2019-2030)*, which includes the ambitious goal of enhancing residents' health literacy to decrease the disease burden and improve national well-being [8].

Given the low-cost and accessibility of ICTs, it is not surprising that policymakers, practitioners, and researchers have focused on electronic health (eHealth) and its applications [9]. Thus, for potential and current individual eHealth users, their eHealth literacy should be considered. eHealth literacy refers to "the ability to seek, find, understand, and appraise health information from electronic sources and apply the knowledge gained to addressing or solving a health problem [10]." Norman and Skinner [10] used the Lily model to describe eHealth literacy in six dimensions: computer literacy, health literacy, traditional literacy, information literacy, science literacy, and media literacy. They also developed the eHealth Literacy Scale (eHEALS) to evaluate each individual's perceptions toward eHealth literacy based on the Lily model [11].

The eHEALS has been shown to be reliable in diverse languages and has been validated in many countries [12-16]. Nevertheless, most of these evaluations were carried out in developed areas, and very few studies in developing areas were reported. eHEALS has also been translated into a traditional Chinese version in Taiwan [17]. For China, the largest developing country with a large rural population, rural-urban disparities exist in terms of cultural adoption, health care resource allocation, and personal health literacy [18,19]. Whether the eHEALS can be used to evaluate Chinese rural residents' eHealth literacy is still unknown. We have two concerns about the existing traditional Chinese version of the eHEALS [17]; it cannot be reliable and valid in mainland China, especially for the Chinese rural population because of two reasons: (1) People living in the Chinese mainland were educated under a simplified Chinese environment. The long-term cultural divide between the Chinese mainland and Taiwan may produce certain semantic differences for specialized vocabularies. The language customs between Chinese mainland and Taiwan are obviously different. For instance, as a typical example of exotic vocabulary, the term "internet" was translated as "  (wang lu)" in traditional Chinese in Taiwan. People who are not familiar with traditional Chinese in mainland China may deem "  " as one type of physical infrastructure rather than the cyber platform. In the modern simplified Chinese context, the "  (lu)" of "  " mostly refers to the physical road. The "internet" should be translated as "  (wang luo)," which semantically emphasizes the network in simplified Chinese. (2) Even in the Chinese mainland, it is still necessary to testify whether the Chinese version of eHEALS is valid and reliable for the rural population. China has been facing an era of internal migration in the recent two decades [20-23]. With the inadequate development and limited work choices in rural areas, rural residents have to migrate to urban areas for better economic benefits, which, in turn, causes a significant issue of rural depopulation [23]. People who cannot migrate to urban areas mostly have a low literacy level and poor health status [24]. One report by the China Internet Network Information Center reveals that rural internet penetration was only 34.0% in mid-2017, while that in urban areas was 69.4% [19]. The perception and knowledge of ICT adoption among rural residents may lag behind those of urban residents. Given the discussion above, the eHEALS must be translated into simplified Chinese, and its psychometric analysis must be performed in the rural population.

The eHEALS was originally proposed to have a one-factor structure [11], which was supported by substantial evidence [12,16,17]. However, recent studies have found that it has a two-factor structure for Italian-speaking people [25], Israeli adults [26], and German adolescents [27]. Additionally, using confirmatory factor analysis, two studies have suggested that eHEALS may have a three-factor structure for baby boomers [28] and outpatients [29]. Although these studies used different statistical strategies, the inconsistent findings imply that the structure of the eHEALS may vary contextually. For instance, Hyde and colleagues indicated that the structure of eHEALS factors differs according to the task complexity [29]. This is true when testing the eHEALS among people living in metropolitan areas, since they have intrinsic modern knowledge distinguishable in terms of complexity levels in use of the eight items of eHEALS. However, the case might be vastly different among people with limited technological and medical literacy in developing areas. The one- or two-factor structure is rational in the latter scenario. Thus, to better understand Chinese rural people's eHealth literacy status, the structure of the eHEALS for the Chinese rural population should be investigated in-depth.

Given the research question proposed above, this study aimed to develop a simplified Chinese version of the eHEALS (C-eHEALS) and evaluate its psychometric properties in a rural population. In this study, both the classical test theory and item response theory methods were adopted based on previous studies' suggestions [25,30].

Methods

Procedure and Participants

In-person interviews on the theme of internet-mobile media usages and health outcomes in rural China were conducted in Chaotian, Sichuan Province, for three weeks in June 2017. Variables included in this study were one part of the entire questionnaire. As of 2017, Chaotian was a poverty-stricken county with 25 towns and a very low level of urbanization. The percentage of rural residents in Chaotian is more than 90% [31]. The quota sampling method was adopted, and each town was assigned 50 quotas considering its individual characteristics such as age, sex, education background, and residential districts. In total, 1250 questionnaires were delivered, and all interviews were conducted by trained local interviewers.

Before the survey, all participants received written information about the study and signed a consent form if they volunteered to participate in the study. When they completed the entire questionnaire, participants were given a small present as compensation. The questionnaires with major illogical, inaccurate, and missing answers accounting for more than 15% of total questions were identified as invalid. In total, 727 rural responses were valid in this survey. After the interviews, researchers randomly selected 30 participants for in-depth interviews about the health-related behaviors' adoption and their influence on participants' daily life. Of all valid responses, 543 participants who identified themselves as internet users were finally included in the analysis of the C-eHEALS.

Measurement

Chinese Version of the eHealth Literacy Scale

This study focuses on the main measurement of the C-eHEALS (Multimedia Appendix 1). Like the original English version of the eHEALS, the C-eHEALS has eight items with response options on a five-point scale, ranging from 1 (strongly disagree) to 5 (strongly agree) [11]. The C-eHEALS was developed following the process of translation and adaptation of the instrument presented by the World Health Organization [32]. First, the eHEALS was translated into simplified Chinese by two bilingual researchers and then reviewed by a bilingual expert panel of four professionals in health communication studies and two rural medical professionals. After the expert panel evaluation, the translated instrument was revised and the complete C-eHEALS was generated. Thereafter, two independent native English translators with no knowledge of the eHEALS translated the C-eHEALS back to English. The resulting items were compared with the original items by the two English translators and the research team to identify possible semantic differences. In addition, the research team compared the simplified Chinese items with the traditional Chinese version of the eHEALS from the previous study [17] to confirm conceptual consistency.

Online Health Information–Seeking Behaviors

We used a multiple-choice question developed from the China Internet Network Information Center [19] to measure individuals' online health information–seeking behaviors as the criterion measure of C-eHEALS. All participants were asked, "In the past 12 months, have you engaged in any of the following behaviors when you accessed the Internet?" We provided the following 11 choices: researched information about hospitals or doctors, researched information about physical exercise, researched information about smoking cessation, researched health or medical information, researched information about drinking cessation, read or shared health information via social media (eg, Weibo and WeChat), researched information about diet, wrote and shared health information via social media (eg, Weibo and WeChat), joined a specific disease internet community, purchased health care products online, and scheduled an appointment online. We added all answers to obtain one indicator—Scope of Online Health Information Seeking Behaviors (SOHISB)—to reflect the diversity in participants' eHealth-related behaviors. Based on the number of behaviors selected, SOHISB ranged from 0 to 11.

Control Variables

For all analyses in this study, several demographic and socioeconomic variables (age, sex, and marital status) were controlled.

Statistical Analysis

First, descriptive statistics, means, SDs, and percentages were calculated for the variables. Second, the C-eHEALS was evaluated according to the classical test theory approach. The reliability of the C-eHEALS was assessed using the Cronbach alpha coefficient (recommended value>.7) [33]. Subsequently, exploratory factor analysis and confirmatory factor analysis were conducted. In the exploratory factor analysis step, the Kaiser-Meyer-Olkin measure of sampling adequacy (recommended value>.6) and the Bartlett test of sphericity (should be statistically significant) [34] were used to test the factorability of C-eHEALS. Principal component analysis was then conducted to examine the latent properties of the eight observed items of C-eHEALS [35] and test whether the structure of the C-eHEALS has a unique pattern or is consistent with the Norman and Skinner one-factor structure [11]. Qualified factors via exploratory factor analysis should account for more than 40% of the total variance with eigenvalues>1 [36,37]. Moreover, a scree plot was used to determine the number of factors to be extracted. In the confirmatory factor analysis, we adopted structural equation modeling to evaluate the structure determined by exploratory factor analysis. The model's goodness of fit was evaluated with the following: the Chi-square value to degrees of freedom ratio (χ^2/df; recommended value<3) [38], comparative fit index (recommended value>.95), Tucker-Lewis index (recommended value >.95), root mean squared error of approximation (recommended value <.06), and standardized root mean squared residual (recommended value<.08) [39].

We thereafter tested other dimensions of the C-eHEALS' psychometric properties using the item response theory approach. In this section, the graded response model [40], which is a generalization of the two-parameter logistic item response model (IRM) for ordinal data, was fit to the C-eHEALS. We chose IRM to help evaluate the C-eHEALS because IRM is more useful than classical test theory in providing information regarding item discriminability and difficulty [25]. Indeed, a recent study suggested that IRM should be used to evaluate the eHEALS properties [25]. Here, the graded response model was adopted over alternative IRMs because each item of the C-eHEALS has five ordered responses. In the graded response model, two types of parameters are generated: discrimination parameter alpha and difficulty or threshold parameter beta [40]. Although alpha indicates how strongly an item relates to a given latent trait theta, beta indicates the level of the latent variable a participant needs to endorse for the next higher response category, with a 50% probability. A larger beta suggests that a higher theta is required for participants to endorse a higher ordered response. Following Baker and Kim's guidelines [41], alpha<.65 indicates low discriminability, .65-1.34 indicates moderate discriminability, and >1.34 indicates high discriminability. Next, item characteristic curves, which present

the probability of participants at a given latent literacy level responding in a particular response category, were estimated for each item [42]. Then, item characteristic curves were transformed into item information curves to demonstrate how much information each item can provide. Thereafter, the test information function, which demonstrates the precision of the entire C-eHEALS along the latent trait continuum, was estimated by summing up all item information curves. In addition, the item characteristic curves were summed, in turn, to obtain the test characteristic curve, which represents the expected score of the C-eHEALS.

Subsequently, following the suggestion of a previous study [14], we also tested the criterion validity of the C-eHEALS via hypothesis testing. We hypothesized that the C-eHEALS score is positively associated with SOHISB among rural participants. Because SOHISB provides count data, Poisson regression was adopted to perform the estimation [43].

Results

Characteristics of the Rural Participants

Details of the characteristics of rural participants are listed in Table 1. Among the 543 participants, the mean age was 40.37 (SD 9.19) years, ranging from 18 to 70 years; men accounted for 58.56% of the participants. Most participants were married (83.43%). As for the highest education level, almost half of the participants listed junior middle school (46.49%), 27.81% listed primary school and below, 16.39% participants listed senior middle school, and 9.21% listed junior college and above. In terms of employment, approximately one-third of the participants (37.57%) had farming jobs, 17.31% had nonfarming jobs within the county, 27.44% had nonfarming jobs outside the county, 7.37% were unemployed, and 10.31% reported other jobs.

Although all samples were recruited in one county via quota sampling, which did not have an ideal representativeness, we compared the internet users in Chaotian and overall China as per a recent report [44]. At the end of 2015, 55.2% Chinese rural internet users were men [44], which was relatively equivalent to the sample in this study. However, the overall Chinese rural internet users showed a higher educational background (20.8% completed primary school and below, 51.9% completed junior middle school, 21.4% completed senior middle school, and 6% completed junior college and above) than the sample in Chaotian. Additionally, regarding the employment status, the overall Chinese rural internet users had fewer people in the farming occupation (15.8%) than the Chaotian sample. The differences between the Chaotian sample and overall Chinese internet users are not surprising as Chaotian, the survey site we chose, was a typical impoverished county when the investigation was conducted.

As for the 11 online health information–seeking behaviors, the least performed behavior was scheduling appointments online (3.31%) and the most was reading or sharing health information via social media (42.91%). As shown in Table 2, the mean of the items in the C-eHEALS ranged from 3.26-3.46 of 5.

Table 1. Characteristics of the rural participants (N=543).

Characteristics	n (%)
Age (years), mean (SD)	40.37 (9.19)
Sex	
Male	318 (58.56)
Female	225 (41.44)
Marital status	
Married	453 (83.43)
Not married	90 (16.57)
Educational background	
Primary school and below	151 (27.81)
Junior middle school	253 (46.59)
Senior middle school	89 (16.39)
Junior college and above	50 (9.21)
Employment status	
Farming	204 (37.57)
Working within the county	94 (17.31)
Working outside the county	149 (27.44)
Nonworking	40 (7.37)
Other	56 (10.31)
Online health information–seeking behavior	
Finding information about hospitals or doctors	96 (17.68)
Finding information about physical exercises	159 (29.28)
Finding information about smoking cessation	36 (6.63)
Finding health or medical information	140 (25.78)
Finding information about drinking cessation	30 (5.52)
Reading or sharing health information via social media (eg, Weibo and Wechat)	233 (42.91)
Finding information about diet	226 (41.62)
Writing and sharing health information via social media (eg, Weibo and Wechat)	200 (36.83)
Attending a specific disease Internet community	62 (11.42)
Purchasing health care products online	38 (7.00)
Online appointment	18 (3.31)

Table 2. Item means for the C-eHEALS in rural participants (N=543).

C-eHEALS items		Mean (SD)
C-eHEALS1	I know what health resources are available on the Internet	3.41 (0.76)
C-eHEALS2	I know where to find helpful health resources on the Internet.	3.32 (0.76)
C-eHEALS3	I know how to find helpful health resources on the Internet.	3.45 (0.73)
C-eHEALS4	I know how to use the Internet to answer my questions about health.	3.33 (0.79)
C-eHEALS5	I know how to use the health information I find on the Internet to help me.	3.46 (0.79)
C-eHEALS6	I have the skills I need to evaluate the health resources I find on the Internet.	3.34 (0.85)
C-eHEALS7	I can tell high quality health resources from low quality health resources on the Internet.	3.33 (0.78)
C-eHEALS8	I feel confident in using information from the Internet to make health decisions.	3.26 (0.87)

Reliability and Exploratory Factor Analysis

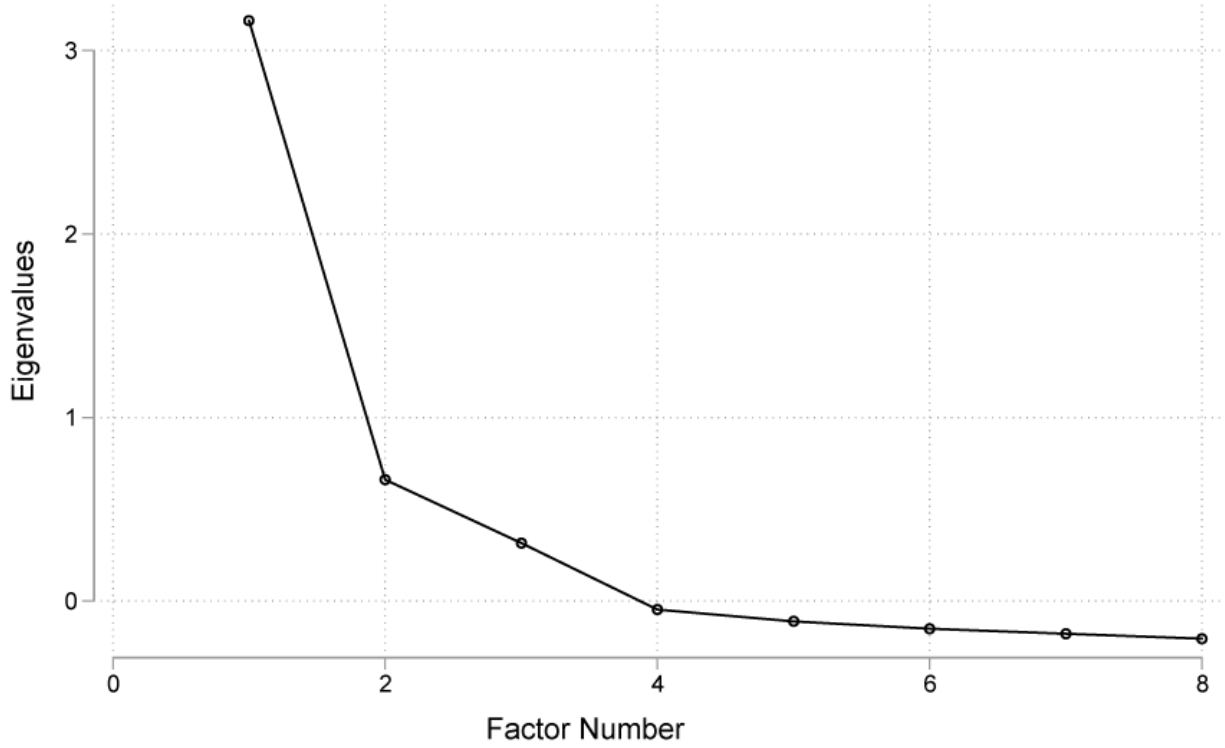
The C-eHEALS had excellent reliability (Cronbach alpha=.834). Both the Kaiser-Meyer-Olkin measure of sampling adequacy (.829; Table 3) and Bartlett test of sphericity (1556.34 (df=28), *P*<.001) showed a good fit to the data, allowing for exploratory factor analysis [34]. Exploratory factor analysis using principal component analysis resulted in a one-factor solution with an initial eigenvalue of 3.159, accounting for 91.8% of the variance (Table 3). The scree plot also showed a one-factor structure (Figure 1). As shown in Table 3, all items loaded above .5, varying from .576 (C-eHEALS1) to .706 (C-eHEALS3). Thus, a single factor was retained.

Table 3. Principal components analysis and Kaiser-Meyer-Olkin test of the C-eHEALS items.

C-eHEALS items	Factor loading	Kaiser-Meyer-Olkin value
C-eHEALS1	0.576	0.820
C-eHEALS2	0.629	0.814
C-eHEALS3	0.706	0.870
C-eHEALS4	0.635	0.820
C-eHEALS5	0.650	0.840
C-eHEALS6	0.582	0.839
C-eHEALS7	0.649	0.785
C-eHEALS8	0.590	0.842
Eigenvalue	3.159	N/A[a]
Cumulative explained variance, %	91.8	N/A
Overall Kaiser-Meyer-Olkin value	N/A	0.829

[a]N/A: not applicable.

Figure 1. Scree plot for Chinese version of the eHealth Literacy Scale.



Confirmatory Factor Analysis

Confirmatory factor analysis was run to verify the one-factor structure obtained from exploratory factor analysis. Results suggested that the C-eHEALS has an excellent one-factor structure (χ^2/df=1.813, comparative fit index=0.993, Tucker-Lewis index=0.985, root mean squared error of

approximation=0.039, standardized root mean squared residual=0.022). Thus, the structure of the C-eHEALS is consistent with the original eHEALS proposed by Norman and Skinner [11] and the traditional Chinese version of the eHEALS [17].

Item Response Theory: Graded Response Models

All eight items of the C-eHEALS were fit to a graded response model. The item parameter estimation and item fit statistics are displayed in Table 4. The discrimination parameters (alpha) ranged from 1.32 to 2.3, indicating that all items discriminated between low and high levels of eHealth literacy well. Only item C-eHEALS6 has a moderate alpha value (1.32), while the other seven items have high discriminability.

Difficulty parameter (beta) estimates indicated that the C-eHEALS is more sensitive at the lower range of latent trait theta because all mean beta ($Beta_M$) values were lower than 0. The beta values of C-eHEALS1 and C-eHEALS5 were unevenly distributed across the trait range, indicating that most participants were unlikely to endorse lower response options. The other six items (C-eHEALS2, C-eHEALS3, C-eHEALS4, C-eHEALS6, C-eHEALS7, and C-eHEALS8) were distributed evenly across the trait range, suggesting that these items differentiate participants from low through high trait levels.

In addition to these results, item characteristic curves are included in Figure 2. These plots show the probability that a participant selects a particular response category at a given level of the latent construct. It was observed that the response categories were distinguishable and monotonically related to the latent trait theta for all items.

Test information function, as reported in Figure 3, reveals that theta values<–3 and >2.5 are poorly represented relative to the rest of the trait range. This is true for seven items (C-eHEALS1, C-eHEALS2, C-eHEALS4, C-eHEALS5, C-eHEALS6, C-eHEALS7, and C-eHEALS8). Only the item information curve of C-eHEALS3 represented a significant fluctuation when theta levels are approximately between 1 and 3. Figure 4 presents the test characteristic curve, which indicates that 95% of randomly selected participants are expected to score between 18.4 and 33.6.

Table 4. Item Response Theory model parameters from C-eHEALS Graded Response Modelsa.

C-eHEALS items	Discrimination			Difficulty				
	alpha	SD	*P* value	$beta_M$[b]	$beta_1$	$beta_2$	$beta_3$	$beta_4$
C-eHEALS1	1.46	0.15	<.001	–0.47	–3.18	–1.96	–0.01	3.26
C-eHEALS2	1.74	0.17	<.001	–0.43	–2.97	–1.77	0.33	2.67
C-eHEALS3	2.31	0.23	<.001	–0.53	–2.86	–1.72	0.06	2.39
C-eHEALS4	1.72	0.16	<.001	–0.51	–3.24	–1.59	0.23	2.57
C-eHEALS5	1.81	0.16	<.001	–0.69	–3.28	–1.70	–0.08	2.31
C-eHEALS6	1.32	0.13	<.001	–0.54	–3.30	–1.78	0.25	2.68
C-eHEALS7	1.55	0.15	<.001	–0.56	–3.58	–1.63	0.25	2.73
C-eHEALS8	1.43	0.14	<.001	–0.38	–2.97	–1.49	0.35	2.59

[a]Discrimination (alpha) refers to an item's ability to discriminate between different latent levels of eHealth literacy (ie, theta). Difficulty parameters (beta) for responses on the 5-point Likert-type scale: 1 (from "strongly disagree" to "disagree"), 2 (from "disagree" to "neutral"), 3 (from "neutral" to "agree"), and 4 (from "agree" to "strongly agree").

[b]$Beta_M$: mean beta.

Figure 2. Item characteristic curves for each Chinese version of the eHealth Literacy Scale. Curves indicate the probability of participants at varying levels of eHealth literacy. C-eHEALS: Chinese version of the eHealth Literacy Scale.

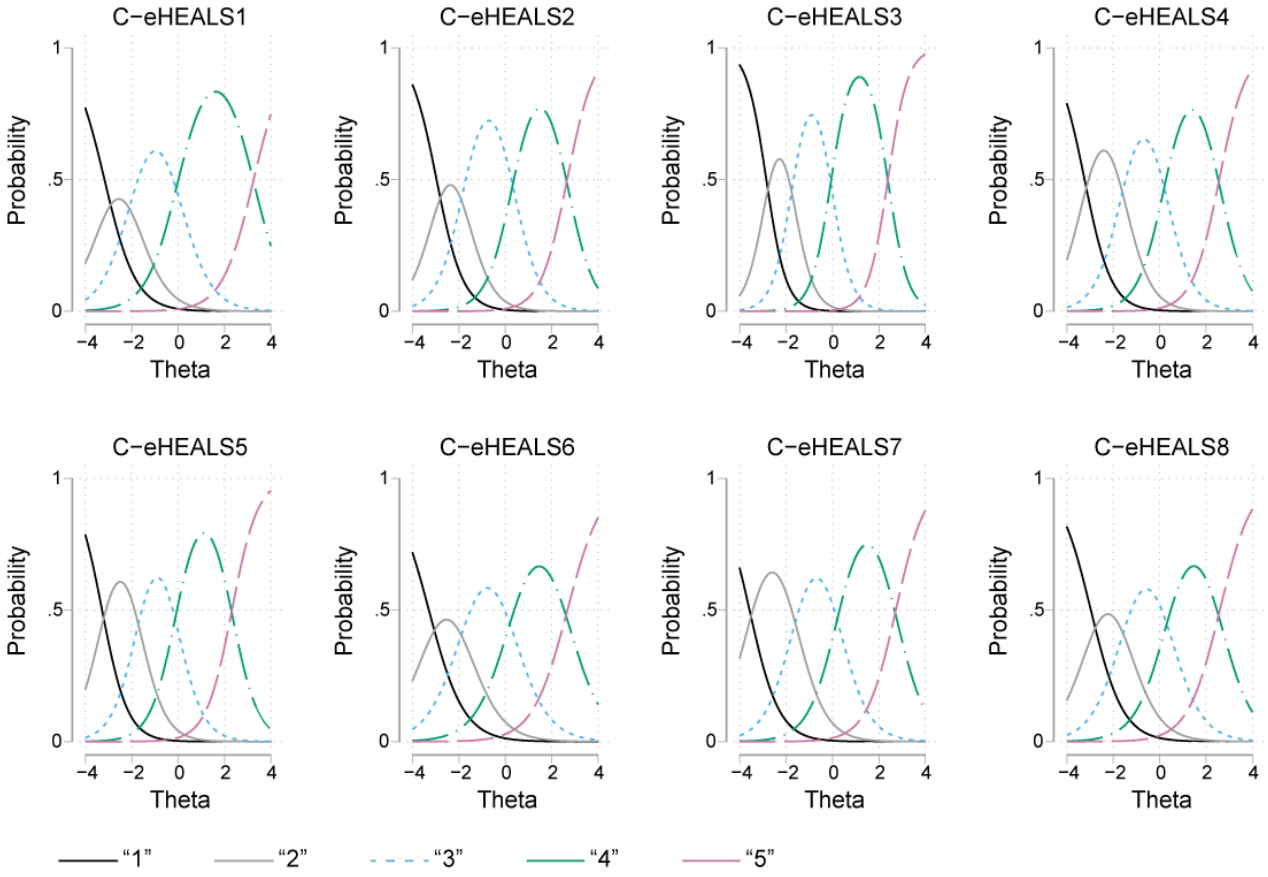


Figure 3. Item information curves and test information function for item characteristic curves. Curves indicate the amount of psychometric information (ie, the reciprocal of the standard error of measurement) provided by the instrument. C-eHEALS: Chinese version of the eHealth Literacy Scale.

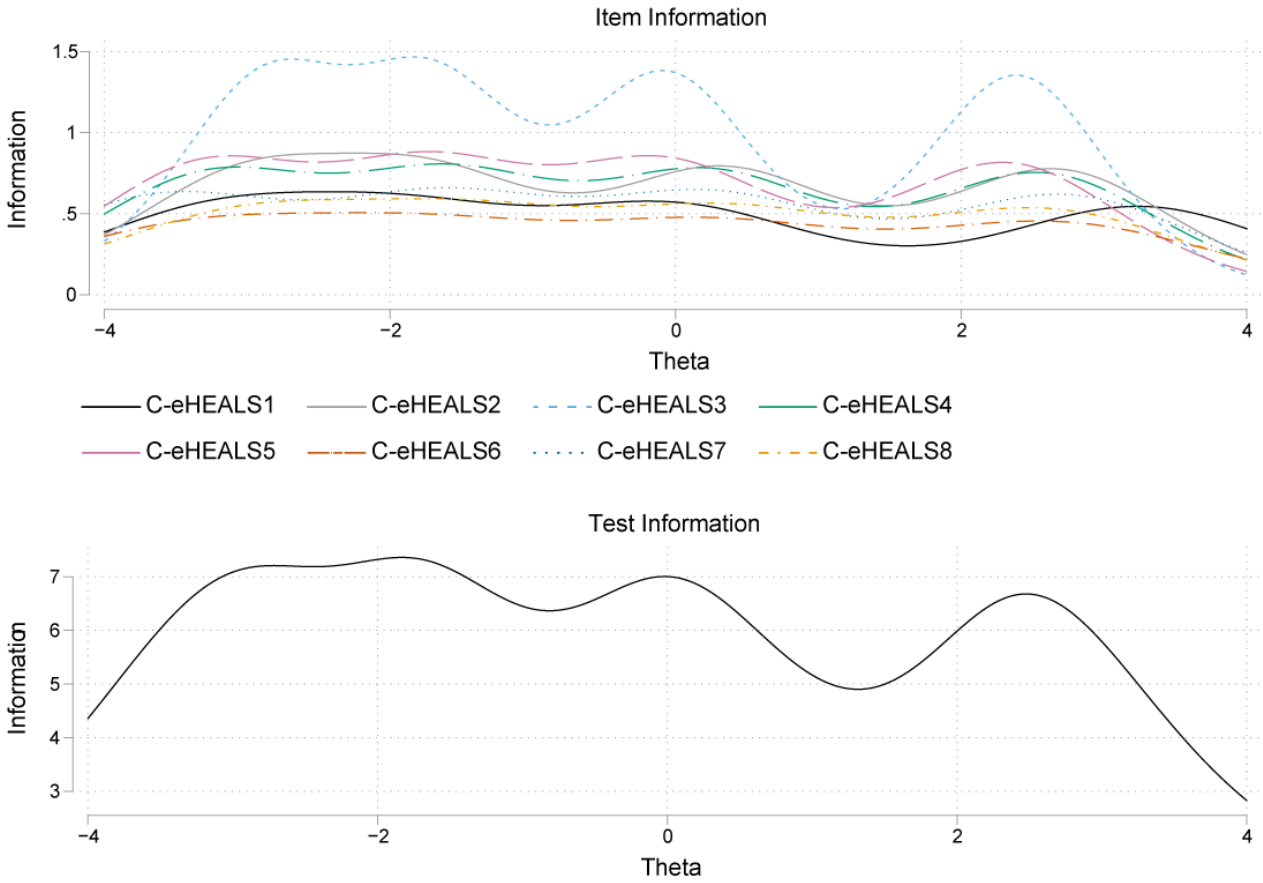
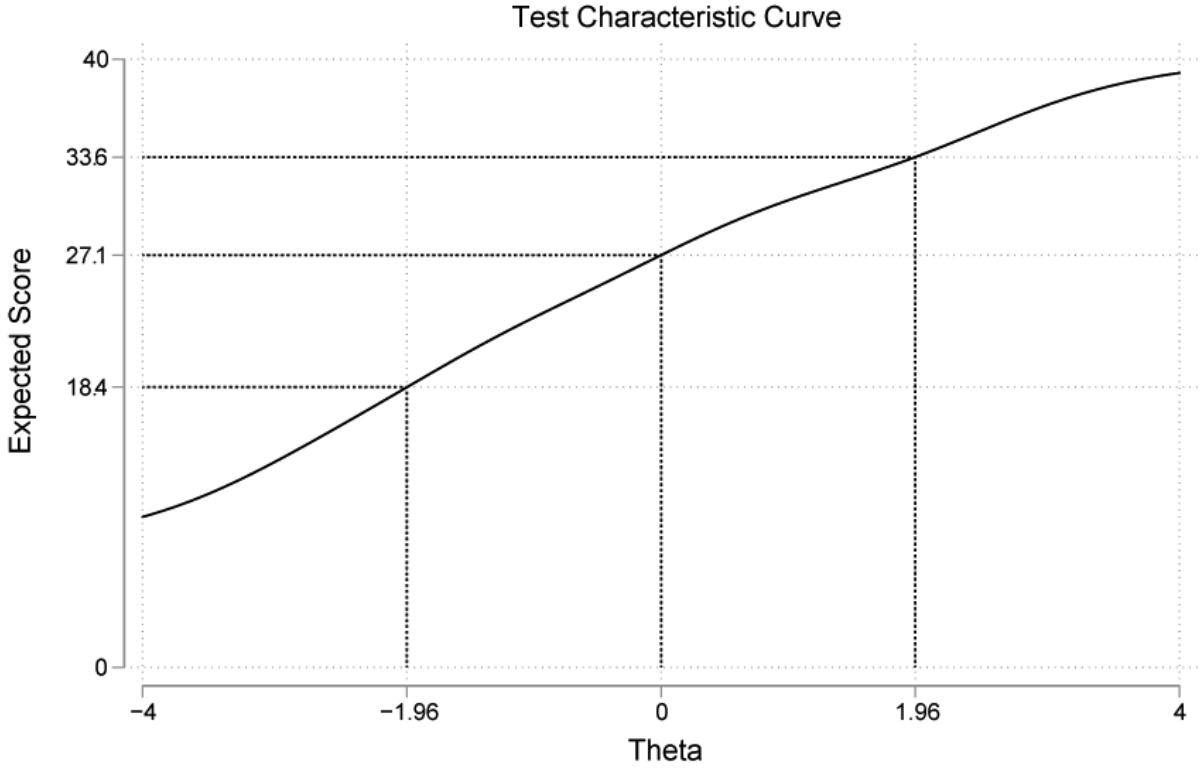


Figure 4. Test characteristic curve for Chinese version of the eHealth Literacy Scale.



Criterion Validity

The results of Poisson regression of the relationship between the C-eHEALS score and SOHISB are displayed in Tables 5 and 6. The odds ratio of the C-eHEALS score is significantly positive, which indicates that participants with higher C-eHEALS scores will have more diverse online health information–seeking behaviors. Hence, the C-eHEALS was shown to have good criterion validity for the Chinese rural population. In addition, coefficients of demographic variables (sex, age, and marital status) do not shown statistical significance, which means the SOHISB does not vary significantly with sex, age, and marital status. However, five dummy variables of socioeconomic status (educational background and employment status) presented significant positive associations with SOHISB. These results are consistent with the knowledge gap hypothesis [45] that implies that individuals will have more possibility to access the information channels to obtain useful information if they are living with higher socioeconomic status. More importantly, even after controlling for all these variables, the C-eHEALS score still plays a positive role in accessing internet technologies to seek health information. The findings in Tables 5 and 6 also suggest that eHealth literacy may have strong practical significance for its implementation to overcome social disparities.

Table 5. Poisson regression results of relationship between C-eHEALS score and SOHISB.

Item	Odds ratio	SD	T value	*P* value	95% CI
C-eHEALS score[a]	1.03	0.07	3.66	<.001	1.01-1.04
Female	1.08	0.07	1.20	.23	0.95-1.22
Age	0.99	0.00	–1.92	.054	0.98-1.00
Married	1.09	0.12	0.80	.42	0.88-1.34
Educational Background (0= primary school and below)					
Junior middle school	1.29	0.10	3.21	.001	1.10-1.51
Senior middle school	1.43	0.15	3.37	.001	1.16-1.76
Junior college and above	1.41	0.17	2.79	.005	1.11-1.80
Employment status (0=farming)					
Working within the county	1.48	0.13	4.64	<.001	1.26-1.75
Working outside the county	1.25	0.10	2.86	.004	1.07-1.46
Not working	1.18	0.15	1.27	.20	0.92-1.51
Other	1.01	0.11	0.07	.95	0.81-1.25
Constant	1.00	0.27	–0.01	.99	0.59-1.70

[a]C-eHEALS: simplified Chinese version of the eHealth Literacy Scale.

Table 6. Poisson regression results.

Model Fit	Value
Observations, n	543
χ^2 (df)	103.2 (11)
Log likelihood	–1062
PR^2	0.0463

Discussion

Principal Findings and Implications

This study investigated the psychometric properties of the simplified C-eHEALS in a Chinese rural population via both classical test theory and item response theory approaches.

Classical test theory analyses demonstrated that the C-eHEALS has good reliability and validity for the rural population in China. The internal consistency of the C-eHEALS was .834, which was comparable to the original eHEALS' alpha value of .88 reported by Norman and Skinner [11]. Exploratory factor analysis results revealed that the C-eHEALS has a one-factor structure, which is consistent with the structures of the original eHEALS [11] and its traditional Chinese version [17]. Surprisingly, the factor accounted for more than 91.8% of the variance, which is much higher than that reported in previous findings [11,14,16,17,25]. Furthermore, referring to previous studies' suggestions [25,29], this one-factor structure also fit well in the confirmatory factor analysis.

Results of the item response theory revealed that response options could differentiate between latent trait levels of all eight C-eHEALS items. The entire instrument provides less information only at extremely low levels (theta<–3) and high

levels (theta>2.5) of the latent trait. These results indicate that C-eHEALS is an excellent measure for capturing participants' ability. Two items (C-eHEALS1 and C-eHEALS5) are more sensitive at the lower range of the latent trait, and the other six items represent excellent discriminability for participants from low through high trait levels.

This study also demonstrated that the C-eHEALS has good criterion validity. A previous study indicated that individuals with a higher level of health literacy will report a larger scope of health information sources [46]. Hence, the diversity of information access channels should be considered as one criterion of better literacy. We also hypothesized that the eHealth literacy score is positively associated with the SOHISB among rural populations. Indeed, controlling for confounding variables, Poisson regression results supported the hypothesis and revealed that rural people's information-seeking behavior could be cultivated with adequate eHealth literacy.

The eHEALS is a validated instrument in diverse language environments [14,16,17,25,27]. In mainland China, it was first introduced in 2013 [47] and has received attention in recent years [48,49]. However, those studies were limited to reporting sophisticated psychometric properties of the eHEALS [47-49]. In addition, rural populations were ignored in previous research. Given the currently targeted poverty alleviation strategies, the campaign launched by the Chinese central government to enhance residents' health literacy status over the next decade [8], and the obvious internet access gap between rural and urban areas [19], both community- and county-level health promotion campaigns should emphasize on health education for rural populations to mitigate large rural-urban disparities.

Health information diffused via the internet should be appropriately evaluated by individual internet users, which can be strengthened by health literacy. For future health literacy-related studies concerning Chinese rural populations, this study provides a useful instrument that can be adopted in survey studies. Moreover, highlighting aspects of health literacy specific to the internet context (eg, practical skills) over other aspects such as the perception description in the eHEALS should be considered for future research.

Findings about the criterion validity also revealed that eHealth literacy is a key element to promote rural residents' access to ICTs for health-related information. Enhancing eHealth literacy might help rural residents overcome their inadequate resource acquisition capacities restricted by local economic recession. It is worth noting that eHealth literacy may also lead to a new digital divide between rural and urban populations. Besides the information-seeking behavior, some other online health-related practices, like health management [4] and mobile app-assisted self-care [50], may present varied implementation practices among people with different socioeconomic statuses. Future studies should design comparisons between populations living in areas with different levels of urbanization.

Limitations

There are two main limitations in this study. First, all participants were recruited from one poverty-stricken county in China via quota sampling, which cannot well represent the diverse situations of China's rural-urban disparities and the entire Chinese rural population. The psychometric properties of the C-eHEALS may vary under different economic development statuses. Hence, future studies aiming to replicate our findings in other samples are highly encouraged. Second, this study had a cross-sectional design, and hence, we were unable to calculate test-retest reliability or predictive validity estimates [51]. Future studies may address this limitation via longitudinal designs.

Conclusions

The C-eHEALS was found to have a robust one-factor structure with excellent discriminability among the Chinese rural population. This scale is helpful for health education practitioners and health professionals to properly measure and understand rural people's eHealth literacy before launching health campaigns. We hope to encourage health researchers who conduct studies in eHealth to carefully investigate policy effects on rural people.

Acknowledgments

This study was supported by the Research Fund of University of Macau (Reference No. MYRG2015-00224-FSS), Key Project of The Key Research Center for Humanities and Social Sciences of Chinese Ministry of Education (Grant No. 16JJD860003), and Fundamental Research Funds for the Central Universities (Grant No. 14370115).

Conflicts of Interest

None declared.

Multimedia Appendix 1
The simplified Chinese version of eHealth Literacy Scale (eHEALS).
[PDF File (Adobe PDF File), 135 KB - jmir_v21i10e15720_app1.pdf]

References

1. CNNIC. 2017. The 39th China Statistical Report on Internet Development URL: http://www.cnnic.net.cn/hlwfzyj/hlwxzbg/hlwtjbg/201701/P020170123364672657408.pdf [accessed 2017-02-24]

2. Hao H. The development of online doctor reviews in China: an analysis of the largest online doctor review website in China. J Med Internet Res 2015 Jun 01;17(6):e134 [FREE Full text] [doi: 10.2196/jmir.4365] [Medline: 26032933]
3. Li Y, Yan X, Song X. Provision of Paid Web-Based Medical Consultation in China: Cross-Sectional Analysis of Data From a Medical Consultation Website. J Med Internet Res 2019 Jun 03;21(6):e12126 [FREE Full text] [doi: 10.2196/12126] [Medline: 31162129]
4. Zhang H, Zhang H, Wang X, Yang Z, Zhao Y. Analysis of Requirements for Developing an mHealth-Based Health Management Platform. JMIR Mhealth Uhealth 2017 Aug 03;5(8):e117 [FREE Full text] [doi: 10.2196/mhealth.5890] [Medline: 28778840]
5. Yingqin Z. Information Culture and Development: Chinese Experience of E-Health. Information Culture and Development: Chinese Experience of E-Health. Proc 38th Annu Hawaii Int Conf Syst Sci Internet IEEE; 2005 Presented at: Proceedings of the 38th Annual Hawaii International Conference on System Sciences; 2005; Big Island, HI, USA p. 153a-1153. [doi: 10.1109/HICSS.2005.315]
6. Fuzhi W, Dan L, Weiwei S, Tingting Y, Dehua H, Wei P, et al. Health Information Literacy and Barriers of Online Health Information Seeking Among Digital Immigrants in Rural China: A Preliminary Survey. SAGE Open 2019 Jun 12;9(2):215824401985694. [doi: 10.1177/2158244019856946]
7. Li AHF. Whom to Trust When Sick? The Wei Zexi Incident, the Chinese Internet and the Healthcare System in China. China Perspect 2016 May:79-83.
8. National Health Commission of the People's Republic China. 2019. Health China Campaign (2019-2030) URL: http://www.nhc.gov.cn/guihuaxxs/s3585u/201907/e9275fb95d5b4295be8308415d4cd1b2.shtml [accessed 2019-07-18]
9. Oh H, Rizo C, Enkin M, Jadad A. What is eHealth (3): a systematic review of published definitions. J Med Internet Res 2005 Feb 24;7(1):e1 [FREE Full text] [doi: 10.2196/jmir.7.1.e1] [Medline: 15829471]
10. Norman CD, Skinner HA. eHealth Literacy: Essential Skills for Consumer Health in a Networked World. J Med Internet Res 2006 Jun 16;8(2):e9 [FREE Full text] [doi: 10.2196/jmir.8.2.e9] [Medline: 16867972]
11. Norman CD, Skinner HA. eHEALS: The eHealth Literacy Scale. J Med Internet Res 2006 Nov 14;8(4):e27 [FREE Full text] [doi: 10.2196/jmir.8.4.e27] [Medline: 17213046]
12. Chung S, Nahm E. Testing reliability and validity of the eHealth Literacy Scale (eHEALS) for older adults recruited online. Comput Inform Nurs 2015 Apr;33(4):150-156 [FREE Full text] [doi: 10.1097/CIN.0000000000000146] [Medline: 25783223]
13. Paige SR, Krieger JL, Stellefson M, Alber JM. eHealth literacy in chronic disease patients: An item response theory analysis of the eHealth literacy scale (eHEALS). Patient Educ Couns 2017 Feb;100(2):320-326. [doi: 10.1016/j.pec.2016.09.008] [Medline: 27658660]
14. Chung S, Park BK, Nahm E. The Korean eHealth Literacy Scale (K-eHEALS): Reliability and Validity Testing in Younger Adults Recruited Online. J Med Internet Res 2018 Apr 20;20(4):e138 [FREE Full text] [doi: 10.2196/jmir.8759] [Medline: 29678800]
15. Mitsutake S, Shibata A, Ishii K, Oka K. Associations of eHealth Literacy With Health Behavior Among Adult Internet Users. J Med Internet Res 2016 Jul 18;18(7):e192 [FREE Full text] [doi: 10.2196/jmir.5413] [Medline: 27432783]
16. van der Vaart R, van Deursen AJ, Drossaert CH, Taal E, van Dijk JA, van de Laar MA. Does the eHealth Literacy Scale (eHEALS) measure what it intends to measure? Validation of a Dutch version of the eHEALS in two adult populations. J Med Internet Res 2011 Nov 09;13(4):e86 [FREE Full text] [doi: 10.2196/jmir.1840] [Medline: 22071338]
17. Koo M, Norman CD, Chang HM, Hsiao-Mei C. Psychometric evaluation of a Chinese version of the eHealth literacy scale (eHEALS) in school age children. Glob J Heal Educ Promot 2012;15:29-36. [doi: 10.5455/gulhane.157832]
18. Zhu Y, Österle A. Rural-urban disparities in unmet long-term care needs in China: The role of the hukou status. Soc Sci Med 2017 Oct;191:30-37. [doi: 10.1016/j.socscimed.2017.08.025] [Medline: 28888126]
19. CNNIC. 2017. The 40th China Statistical Report on Internet Development URL: http://www.cnnic.net.cn/hlwfzyj/hlwxzbg/hlwtjbg/201708/P020170807351923262153.pdf [accessed 2018-03-12]
20. Liang Z, Por Chen Y, Gu Y. Rural Industrialisation and Internal Migration in China. Urban Studies 2016 Jul 02;39(12):2175-2187. [doi: 10.1080/0042098022000033926]
21. Fan CC. Interprovincial Migration, Population Redistribution, and Regional Development in China: 1990 and 2000 Census Comparisons. The Professional Geographer 2005 May;57(2):295-311. [doi: 10.1111/j.0033-0124.2005.00479.x]
22. Li Y, Chen C, Wang Y, Liu Y. Urban-rural transformation and farmland conversion in China: The application of the environmental Kuznets Curve. Journal of Rural Studies 2014 Oct;36:311-317. [doi: 10.1016/j.jrurstud.2014.10.005]
23. Li Y, Westlund H, Zheng X, Liu Y. Bottom-up initiatives and revival in the face of rural decline: Case studies from China and Sweden. Journal of Rural Studies 2016 Oct;47:506-513. [doi: 10.1016/j.jrurstud.2016.07.004]
24. Lu Y, Qin L. Healthy migrant and salmon bias hypotheses: a study of health and internal migration in China. Soc Sci Med 2014 Feb;102:41-48. [doi: 10.1016/j.socscimed.2013.11.040] [Medline: 24565140]
25. Diviani N, Dima AL, Schulz PJ. A Psychometric Analysis of the Italian Version of the eHealth Literacy Scale Using Item Response and Classical Test Theory Methods. J Med Internet Res 2017 Apr 11;19(4):e114 [FREE Full text] [doi: 10.2196/jmir.6749] [Medline: 28400356]
26. Neter E, Brainin E, Baron-Epel O. The dimensionality of health literacy and eHealth literacy. European Health Psychologist 2015;17(6):275-280.

27. Soellner R, Huber S, Reder M. The Concept of eHealth Literacy and Its Measurement. Journal of Media Psychology 2014 Jan;26(1):29-38. [doi: 10.1027/1864-1105/a000104]
28. Sudbury-Riley L, FitzPatrick M, Schulz PJ. Exploring the Measurement Properties of the eHealth Literacy Scale (eHEALS) Among Baby Boomers: A Multinational Test of Measurement Invariance. J Med Internet Res 2017 Feb 27;19(2):e53 [FREE Full text] [doi: 10.2196/jmir.5998] [Medline: 28242590]
29. Hyde LL, Boyes AW, Evans T, Mackenzie LJ, Sanson-Fisher R. Three-Factor Structure of the eHealth Literacy Scale Among Magnetic Resonance Imaging and Computed Tomography Outpatients: A Confirmatory Factor Analysis. JMIR Hum Factors 2018 Feb 19;5(1):e6 [FREE Full text] [doi: 10.2196/humanfactors.9039] [Medline: 29459356]
30. Prisciandaro JJ, Tolliver BK. An item response theory evaluation of the young mania rating scale and the montgomery-asberg depression rating scale in the systematic treatment enhancement program for bipolar disorder (STEP-BD). J Affect Disord 2016 Nov 15;205:73-80 [FREE Full text] [doi: 10.1016/j.jad.2016.06.062] [Medline: 27420134]
31. GYCT. 2015. Statistical Bulletin on National Economic and Social Development in Chaotian (2014) URL: http://www.gyct.com.cn/info/1487/27774.htm [accessed 2019-07-17]
32. World Health Organisation. 2010. Process of translation and adaptation of instruments URL: http://www.who.int/substance_abuse/research_tools/translation/en/ [accessed 2019-07-17]
33. Furnham A. A handbook of test construction, introduction to psychometric design. Personality and Individual Differences 1987 Jan;8(5):773. [doi: 10.1016/0191-8869(87)90086-9]
34. Dziuban CD, Shirkey EC. When is a correlation matrix appropriate for factor analysis? Some decision rules. Psychological Bulletin 1974;81(6):358-361. [doi: 10.1037/h0036316]
35. Park HS, Dailey R, Lemus D. The Use of Exploratory Factor Analysis and Principal Components Analysis in Communication Research. Human Comm Res 2002 Oct;28(4):562-577. [doi: 10.1111/j.1468-2958.2002.tb00824.x]
36. Glorfeld LW. An Improvement on Horn's Parallel Analysis Methodology for Selecting the Correct Number of Factors to Retain. Educational and Psychological Measurement 2016 Jul 02;55(3):377-393. [doi: 10.1177/0013164495055003002]
37. Hayton JC, Allen DG, Scarpello V. Factor Retention Decisions in Exploratory Factor Analysis: a Tutorial on Parallel Analysis. Organizational Research Methods 2016 Jun 29;7(2):191-205. [doi: 10.1177/1094428104263675]
38. Hoe S. Issues and procedures in adopting structural equation modeling technique. Journal of Applied Quantitative Methods 2008;3(1):76-83.
39. Hu L, Bentler P. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Structural Equation Modeling: A Multidisciplinary Journal 1999 Jan;6(1):1-55. [doi: 10.1080/10705519909540118] [Medline: 3384345]
40. Samejima F. Estimation of latent ability using a response pattern of graded scores. ETS Research Bulletin Series 2014 Aug 08;1968(1):i-169. [doi: 10.1002/j.2333-8504.1968.tb00153.x]
41. Frank B, Kim SH. Item response theory: Parameter estimation techniques. New York, NY: Marcel Dekker; 2004.
42. Mielenz TJ, Callahan LF, Edwards MC. Item response theory analysis of Centers for Disease Control and Prevention Health-Related Quality of Life (CDC HRQOL) items in adults with arthritis. Health Qual Life Outcomes 2016 Mar 12;14(1):43 [FREE Full text] [doi: 10.1186/s12955-016-0444-4] [Medline: 26969513]
43. Kubinyi E, Miklósi A, Topál J, Csányi V. Social mimetic behaviour and social anticipation in dogs: preliminary results. Anim Cogn 2003 Mar;6(1):57-63. [doi: 10.1007/s10071-003-0163-1] [Medline: 12658535]
44. CNNIC. 2016. Statistical Report on Rural Internet Development - 2015 URL: https://www.cnnic.net.cn/hlwfzyj/hlwxzbg/ncbg/201608/P020170907348967498375.pdf [accessed 2018-03-12]
45. Donohue G, Tichenor P, Olien C. Mass Media and the Knowledge Gap. Communication Research 2016 Jun 30;2(1):3-23. [doi: 10.1177/009365027500200101]
46. Kim Y, Lim JY, Park K. Effects of Health Literacy and Social Capital on Health Information Behavior. J Health Commun 2015;20(9):1084-1094. [doi: 10.1080/10810730.2015.1018636] [Medline: 26166008]
47. Guo S, Yu X, Sun Y, Nie D, Wang L. Adaptation and evaluation of Chinese version of eHEALS and its usage among senior high school students. Chinese Journal of Health Education 2013;29(2):106-108.
48. Zhao Y, Chen H, Gong H, Wu Y. Analysis on status and influencing factor of e-health literacy among outpatients. Modern Preventive Medicine 2019;46(6):1073-1073.
49. Zhou H, Zheng A. Current situation and influencing factors of e-health literacy among community older adults. Journal of Nanjing Medical University (Social Sciences) 2018;18(6):455-458.
50. Ploderer B, Brown R, Seng LSD, Lazzarini PA, van Netten JJ. Promoting Self-Care of Diabetic Foot Ulcers Through a Mobile Phone App: User-Centered Design and Evaluation. JMIR Diabetes 2018 Oct 10;3(4):e10105 [FREE Full text] [doi: 10.2196/10105] [Medline: 30305266]
51. Weir JP. Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. J Strength Cond Res 2005 Feb;19(1):231-240. [doi: 10.1519/15184.1] [Medline: 15705040]

Abbreviations

C-eHEALS: simplified Chinese version of the eHealth Literacy Scale

eHEALS: eHealth Literacy Scale
eHealth: electronic health
ICTs: information communication technologies
IRM: item response model
SOHISB: Scope of Online Health Information Seeking Behaviors

Edited by G Eysenbach; submitted 01.08.19; peer-reviewed by E Neter, J Papadakos, G Bravo, O Baron-Epel, C Bakker; comments to author 28.08.19; revised version received 06.09.19; accepted 23.09.19; published 22.10.19.

Please cite as:
Ma Z, Wu M
The Psychometric Properties of the Chinese eHealth Literacy Scale (C-eHEALS) in a Chinese Rural Population: Cross-Sectional Validation Study
J Med Internet Res 2019;21(10):e15720
URL: http://www.jmir.org/2019/10/e15720/
doi:10.2196/15720
PMID:31642811

Original Paper

Results of MyPlan 2.0 on Physical Activity in Older Belgian Adults: Randomized Controlled Trial

Delfien Van Dyck[1,2], PhD; Karel Herman[1], MSc; Louise Poppe[1,2], MSc; Geert Crombez[3], PhD; Ilse De Bourdeaudhuij[1,4], PhD; Freja Gheysen[1,5], PhD

[1]Department of Movement and Sports Sciences, Faculty of Medicine and Health Sciences, Ghent University, Ghent, Belgium

[2]Research Foundation – Flanders, Brussels, Belgium

[3]Department of Experimental Clinical and Health Psychology, Faculty of Psychology and Educational Sciences, Ghent University, Ghent, Belgium

[4]Department of Educational Policy, Ghent University, Ghent, Belgium

[5]Department of Research and Valorisation, Vives University of Applied Sciences, Kortrijk, Belgium

Corresponding Author:
Delfien Van Dyck, PhD
Department of Movement and Sports Sciences
Faculty of Medicine and Health Sciences
Ghent University
Watersportlaan 2
Ghent, 9000
Belgium
Phone: 32 92646323
Email: delfien.vandyck@ugent.be

Abstract

Background: The beneficial effects of physical activity (PA) for older adults are well known. However, few older adults reach the health guideline of 150 min per week of moderate-to-vigorous PA (MVPA). Electronic health (eHealth) interventions are effective in increasing PA levels in older adults in the short term but, rarely, intermediate-term effects after a period without the support of a website or an app have been examined. Furthermore, current theory-based interventions focus mainly on preintentional determinants, although postintentional determinants should also be included to increase the likelihood of successful behavior change.

Objective: This study aimed to investigate the effect of the theory-based eHealth intervention, MyPlan 2.0, focusing on pre- and postintentional determinants on both accelerometer-based and self-reported PA levels in older Belgian adults in the short and intermediate term.

Methods: This study was a randomized controlled trial with three data collection points: baseline (N=72), post (five weeks after baseline; N=65), and follow-up (three months after baseline; N=65). The study took place in Ghent, and older adults (aged ≥65 years) were recruited through a combination of random and convenience sampling. At all the time points, participants were visited by the research team. Self-reported domain-specific PA was assessed using the International Physical Activity Questionnaire, and accelerometers were used to objectively assess PA. Participants in the intervention group got access to the eHealth intervention, MyPlan 2.0, and used it independently for five consecutive weeks after baseline. MyPlan 2.0 was based on the self-regulatory theory and focused on both pre- and postintentional processes to increase PA. Multilevel mixed-models repeated measures analyses were performed in R (R Foundation for Statistical Computing).

Results: Significant (borderline) positive intervention effects were found for accelerometer-based MVPA (baseline−follow-up: intervention group +5 min per day and control group −5 min per day; *P*=.07) and for accelerometer-based total PA (baseline−post: intervention group +20 min per day and control group −24 min per day; *P*=.05). MyPlan 2.0 was also effective in increasing self-reported PA, mainly in the intermediate term. A positive intermediate-term intervention effect was found for leisure-time vigorous PA (*P*=.02), moderate household-related PA (*P*=.01), and moderate PA in the garden (*P*=.04). Negative intermediate-term intervention effects were found for leisure-time moderate PA (*P*=.01) and cycling for transport (*P*=.07).

Conclusions: The findings suggest that theory-based eHealth interventions focusing on pre- and postintentional determinants have the potential for behavior change in older adults. If future studies including larger samples and long-term follow-up can

confirm and clarify these findings, researchers and practitioners should be encouraged to use a self-regulation perspective for eHealth intervention development.

Trial Registration: Clinicaltrials.gov NCT03194334; https://clinicaltrials.gov/ct2/show/NCT03783611.

(J Med Internet Res 2019;21(10):e13219) doi:10.2196/13219

KEYWORDS

self-regulation; exercise; elderly; eHealth

Introduction

Background

The beneficial effects of physical activity (PA) for older adults (aged ≥65 years) are well known. PA reduces the risk of developing common chronic diseases, such as type 2 diabetes, cardiovascular diseases, and hypertension. In addition, PA has a positive effect on overall physical and mental functioning and on morbidity and mortality rates [1-5]. However, many older adults are not sufficiently active [3]. The World Health Organization (WHO) states that "older adults should do at least 150 min of moderate-intensity aerobic PA or 75 min of vigorous-intensity aerobic PA throughout the week, or a combination of both." [6]. However, depending on the country, 60% to 70% of older adults in Western countries do not reach the PA health guideline [7]. Similarly, only 31% of older Belgian adults aged between 65 and 74 years are sufficiently physically active [8]. In those aged 75 years and older, this is only 12% [8]. Given these low levels of PA, it is necessary to develop effective interventions for this particular age group [9].

Overall, health behavior interventions are often not theory based. Nonetheless, it has been shown that the use of a theoretical framework for intervention development enhances the effectiveness of an intervention [10,11]. For example, the theoretical framework of self-regulation is useful for intervention development [12]. Self-regulation is defined as "a goal-guidance process aimed at the attainment and maintenance of personal goals" [12]. The process of behavior change can be divided in a pre- and a postintentional phase. In the preintentional phase, an individual acknowledges a problem (eg, the lack of PA) and develops intentions to solve this problem. In the postintentional phase, an individual sets goals and makes action plans to achieve them. In the past, interventions to increase PA levels in older adults that made use of a theoretical framework primarily targeted preintentional determinants (eg, attitude, self-efficacy, and expected outcomes) of PA [13]. However, changing these determinants does not necessarily imply that people will change their actual behavior. This is the so-called *intention behavior gap* [14]. To achieve actual behavior change, postintentional determinants (eg, making action plans and engaging in goal pursuit and goal adaptation) must also be integrated in an intervention. By focusing on the whole process of behavior change, the likelihood of successful behavior change increases.

In the last decade, researchers started using mobile apps and websites to promote PA and well-being in different age groups [15,16]. One of the major advantages of this evolution is the increasing accessibility of health care. In addition, face-to-face contact is no longer needed, and tailored interventions can be executed at home [17]. Furthermore, delivering electronic health (eHealth) interventions is less expensive than providing traditional interventions [18]. Research also indicated that eHealth interventions are suitable for older adults [19]. In 2015, The Federal Public Service of economy of Belgium reported that approximately 73% of adults aged 65 to 74 years used the internet on a daily basis. As this percentage is still increasing, eHealth interventions become more and more appropriate to promote PA in older adults.

Previous studies already showed that a tailored eHealth intervention, based on the self-regulation theory, could increase PA in (older) adults [15,16,20]. Degroote et al [20] and Plaete et al [16] showed that *MyPlan 1.0*, a website based on the self-regulation theory and the Health Action Process Approach (a specific model of self-regulation [21]), was effective in increasing PA levels in adults after a month of intervention. Similar effects were found in older adults [15].

Despite the promising results of the previous *MyPlan 1.0* intervention studies, several research questions remain unanswered. In the 3 studies mentioned above [15,16,20], assessments took place a week and a month after the start of the intervention, that is, after a period of continuous website support. On the basis of this protocol, it is impossible to determine whether these effects last for a longer period, especially when support from the website is no longer being provided. Overall, such evidence is still lacking [22]. As it is important to maintain a physically active lifestyle [23], this study will focus on the effects of the eHealth website, *MyPlan 2.0*, on PA levels in older adults after a period of 2 months without website support.

Objective

Furthermore, previous studies mainly used self-reported PA data that are known to be subject to recall bias and over-reporting [24]. To overcome this problem, this study combines self-reported and objective methods to assess PA. Consequently, the aim of this study was to examine the short- and intermediate-term effects of the *MyPlan 2.0* eHealth intervention on objectively measured and self-reported PA levels in older adults. It was expected that self-reported and objectively measured PA levels would increase in the intervention group immediately after using the website for 5 weeks (short-term effects), compared with the control group. As the intervention was based on self-regulation, that is, guiding individuals gradually toward their goals [12,16], it was also expected that the short-term effects would be maintained in the intermediate term, after a period without website support.

Methods

Study Design

This study was a parallel randomized controlled trial (1:1 allocation) using random sampling in combination with convenience sampling.

Research Site

The study took place in Ghent and its suburbs. Ghent is the second largest city in Flanders, which is the Dutch-speaking part of Belgium. It has approximately 260,000 inhabitants.

Procedure

First, the Public Service of Ghent provided names and addresses of 1000 randomly selected adults aged between 65 and 80 years. Second, the research team randomly sent 500 invitation letters to participate in the *MyPlan 2.0* intervention and 500 invitation letters to be part of a control group receiving no intervention. As the response rate was very low, the research team additionally recruited participants by handing out flyers in local service centers (ie, convenience sampling).

The inclusion criteria for this study were the following: being aged 65 to 80 years, retired, able to walk 100 m without any help (ie, devices or help from persons), Dutch-speaking, and have an email address. The email address was needed for logging in to the website and for sending the weekly reminders to visit the website. Eligible participants were asked to confirm their participation by email or phone. Afterward, participants received an email with extra information about the study.

The study comprised 6 appointments in person (Figure 1). Data were collected from November 2016 to June 2017. During the first appointment (baseline data collection), all participants signed the informed consent, filled out the long International Physical Activity Questionnaire (IPAQ, interview version) and a demographic questionnaire, and received an accelerometer (Figure 2). At least one week later (appointment 2), the accelerometer was recollected and participants of the intervention group were invited to use the *MyPlan 2.0* intervention for 5 consecutive weeks (ie, 5 website visits). The control group did not get access to the website. After 5 weeks, when the intervention group completed the *MyPlan 2.0* intervention, the post data collection took place: participants were interviewed (long IPAQ) and asked to wear the accelerometer for the second time (appointment 3). A week later, the accelerometer was recollected (appointment 4). A total of 3 months after baseline, follow-up measurements were conducted. The participants wore the accelerometer and filled out the IPAQ (interview) for the last time (appointment 5). A week later, during the final appointment, the accelerometer was recollected. The study protocol was approved by the Ethics Committee of the Ghent University Hospital (project number 2015/1502).

Figure 1. Study design of MyPlan 2.0. IPAQ: International Physical Activity Questionnaire.

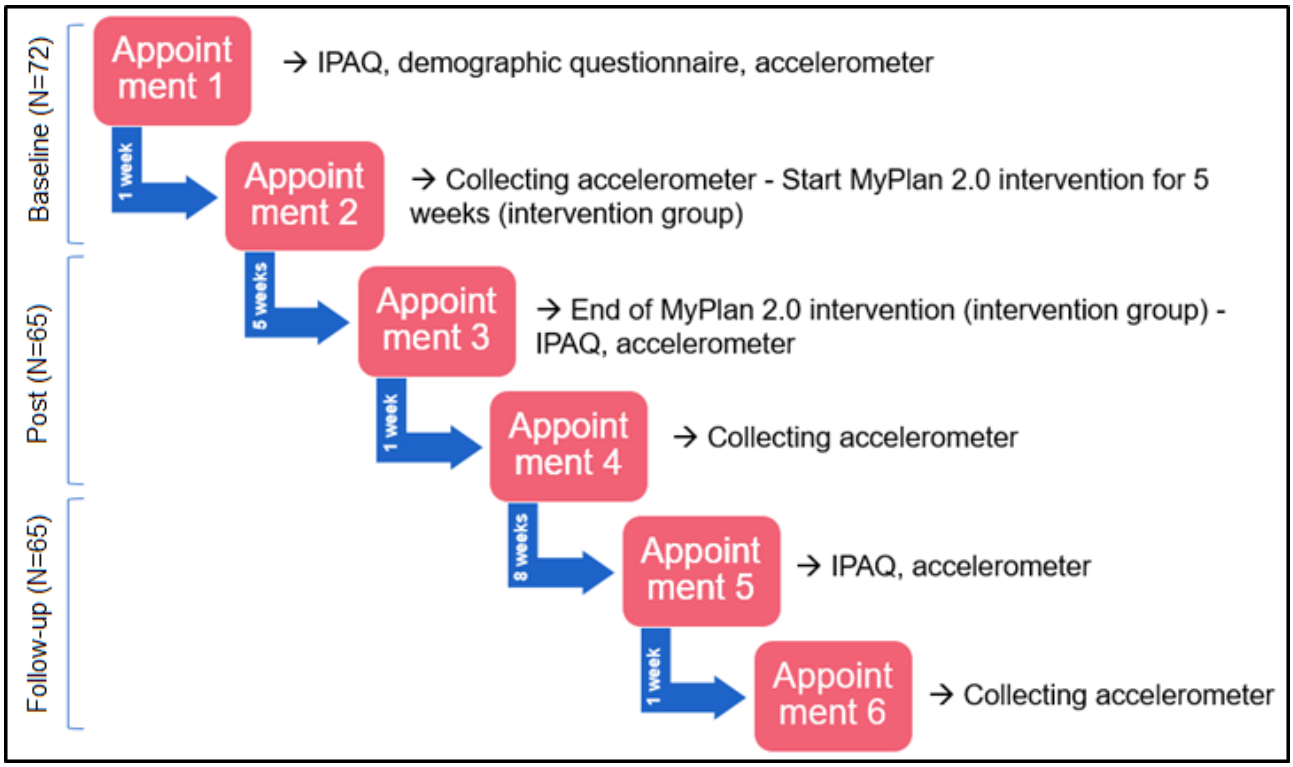
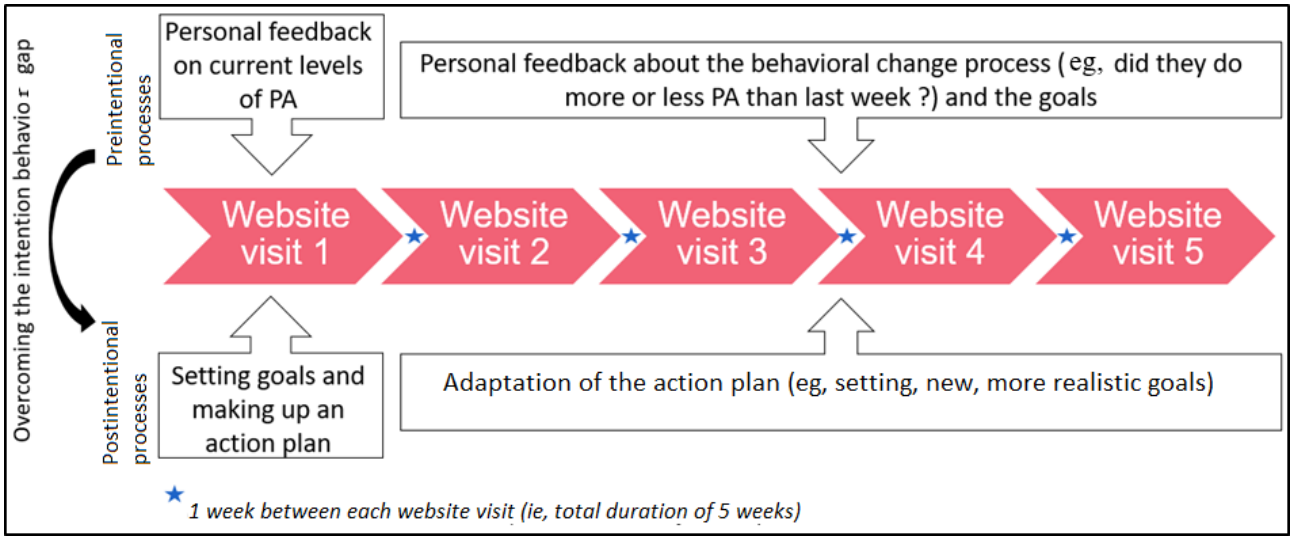


Figure 2. Overview of the electronic health intervention MyPlan 2.0. PA: physical activity.



Intervention

In this study, *MyPlan 2.0*, an improved version of *MyPlan 1.0* [20,21], was used. *MyPlan 1.0* was mainly theory based, whereas *MyPlan 2.0* is theory- and user-based. Several qualitative studies were performed to optimally adapt the intervention to the users' needs. For example, users of *MyPlan 1.0* indicated that they felt demotivated by the extensive questionnaires they had to complete to receive tailored feedback and stated that they did not understand why creating coping plans would help alter their behavior [25]. In *MyPlan 2.0*, these questionnaires were significantly shortened and rationales for the implemented behavior change techniques were added. Moreover, Vandelanotte et al [26] showed that interventions with minimum 5 contact moments (eg, appointments, Web modules, and emails) were more successful. Therefore, *MyPlan 2.0* comprised 5 website visits in contrast to the 3 obligatory website visits of *MyPlan 1.0*. The first website visit (for details see below) contained pre- and postintentional processes. The following 4 website visits mainly contained postintentional processes. Participants could independently use the website, without researcher involvement.

During the first website visit, participants had to complete a short PA questionnaire and based on the answers, they received computer-tailored or personalized feedback. By doing so, the preintentional processes were targeted (Figure 2). This personalized feedback was based on a comparison of the users' PA levels with the health guidelines of 150 min per week of moderate-to-vigorous PA (MVPA) [6]. To increase knowledge, users had the option to complete a quiz about PA and its beneficial effects. As shown in Figure 2, postintentional processes were targeted by asking the participants to make an action plan. By doing so, the gap between intentions and behavior was bridged. Participants were asked *what* they wanted to do (eg, being more active by cycling during leisure), *when* (eg, every Sunday morning), *where* (eg, in the streets nearby), and *for how long* (eg, 60 min) they were planning to do the activity. After providing answers to these questions, participants could identify difficult situations and possible barriers (ie, coping planning) while pursuing their goals, using a predefined list of situations and barriers. Depending on which barriers they selected, specific solutions were given, and participants could choose which ones they considered most appropriate and applicable. At the end of this first website visit, users could indicate how they wanted to self-monitor their behavior (eg, using an agenda), and they could read more information about how to receive support toward PA from their social environment. Finally, the personal action plan could be printed weekly (optional). Multimedia Appendix 1 provides screenshots from the website and links these to the self-regulation techniques that were used.

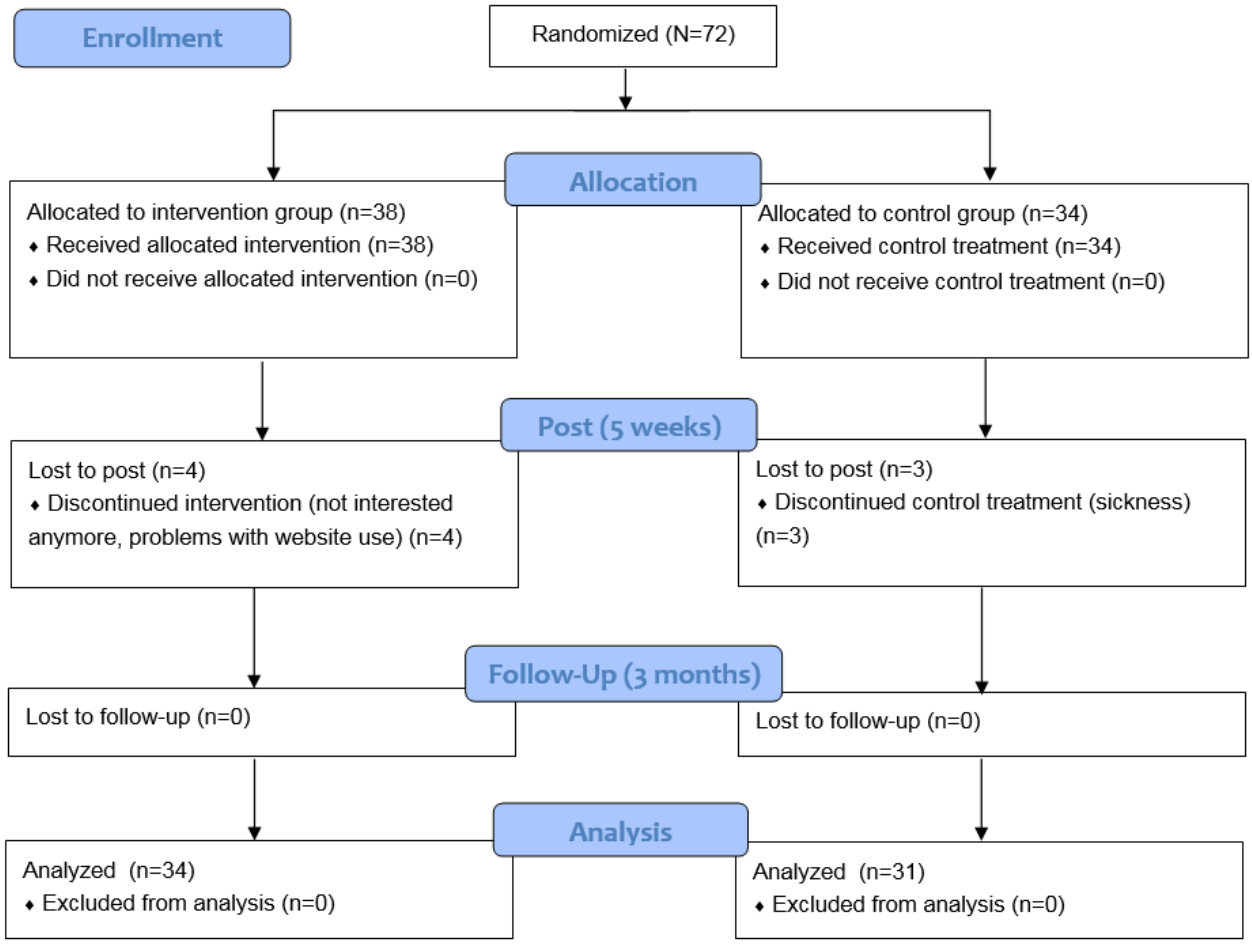
A week after finishing the first website visit, participants received an email to revisit the website. During this second visit, they received feedback about their behavioral change process and goals (eg, *did you reach your goal or not?*). Afterward, participants had the possibility to adapt their action plan (eg, setting new, more realistic goals) and reconsider coping plans based on the barriers they experienced while pursuing their goals. Furthermore, participants could optionally read tips on how to increase PA.

Website visits 3, 4, and 5 were respectively activated 1 week after the previous visit. Again, participants were reminded by email. These 3 last visits were identical to the second visit (reviewing the action and coping plans). If participants did not revisit the website after 1 week, the research team phoned them reminding them to revisit the website. Figure 2 provides an overview of the intervention.

Participants

At baseline, the total sample comprised 72 older adults, 38 in the intervention group and 34 in the control group. Between pre and post measurements, 7 people dropped out. Of them, 3 were part of the control group and 4 of the intervention group. Reasons for dropping out were as follows: no longer interested in the intervention (n=2), sickness (n=3), and problems with using the website (n=2). There was no dropout between post and follow-up measurements (Figure 3).

Figure 3. Participant flow diagram.



Instruments and Materials

The following sociodemographic variables were assessed at baseline: age, gender, height and weight, marital status (married, widowed, divorced, cohabiting, and living alone), and highest degree of education (primary school, secondary school, college, and university).

Self-reported PA was assessed at baseline, post, and follow-up, using the long Dutch IPAQ interview version (usual week version). This questionnaire assesses the frequency and duration of walking, cycling, moderate-intensity PA, and vigorous-intensity PA in 4 domains: (voluntary) work, transport, leisure, and household (home and garden). The IPAQ has good reliability (intraclass range from 0.46-0.96), and the criterion validity is fair-to-moderate with Spearman rho ranging from 0.30 to 0.37 [27,28]. As the IPAQ has a tendency of over-reporting, all data were truncated according to the official IPAQ guidelines [29].

Objective PA was assessed at baseline, post, and follow-up, using an ActiGraph GT3X+ accelerometer. Participants wore the accelerometer for 7 days on the right hip. They were asked to wear it during waking hours but not when swimming, showering, or practicing a contact sport. The accelerometers were initialized and processed using Actilife 6.13.3. Valid wear time was set as at least four days with at least ten hours of wear time. Non–wear time was defined as ≥60 min of consecutive zeros. The epoch was set at 60 seconds, and the cut point used to determine MVPA was set at 1952 counts per minute (cpm) [30]. The cut point for light-intensity activity was set at 100 to 1951 cpm. The level of total PA was calculated by adding up light-intensity PA and MVPA.

Statistical Analyses

Baseline sociodemographic characteristics of the intervention and control group were compared using independent sample *t* tests (continuous variables) and chi-square tests (categorical variables) in SPSS 25.0. To evaluate the intervention effects on accelerometer-assessed and self-reported PA, multilevel mixed-models repeated measures analyses were performed in R (package lme4) [31]. Multilevel modeling (2-level: measurement–participant) was applied to take into account the clustering of the 3 measurements (pre–post–follow-up) in participants. On the basis of the recommendations of Chakraborty and Gu [32], no ad hoc data imputation was applied. As almost all PA variables, except for accelerometer-based total PA, were positively skewed, square-root transformations were applied to improve normality. To increase the comprehensibility of the tables, raw descriptive data have been reported, although analyses were conducted using the square-root transformed data. For each PA variable (2 accelerometer-based and 11 self-reported PA variables), a

separate regression model was fitted. The reported beta value for the interaction effect between *time* and *condition* can be interpreted as the difference in change in outcome between pre- and posttest, pre- and follow-up test, and post- and follow-up test according to the condition to which participants belong (intervention vs control condition). Statistical significance was set at $P<.05$ but because of the small sample size, borderline significant results ($P<.10$) were also reported.

Results

Participants

Baseline descriptive statistics are shown in Table 1. At baseline, 72 older adults (38 in intervention group and 34 in control group) participated in this study, 51% (37/72) were male. The participants' mean age was 70.9 (SD 4.1) years, and mean body mass index was 26.4 (SD 4.2) g/m^2. In total, 64% (46/72) of all participants were married, and 47% (34/72) had a college or university degree. There were no significant baseline differences in sociodemographic characteristics between the intervention and the control group. Consequently, no covariates were included in further analyses.

Table 1. Sociodemographic characteristics at baseline.

Sociodemographic characteristics	Total sample (N=72)	Control group (n=34)	Intervention group (n=38)	X^2 value (*df*)
Age (years), mean (SD)	70.9 (4.1)	70.9 (4.1)	70.8 (4.1)	0.1 (*70*)[a]
Gender, n (%)				
Male	37 (51)	19 (56)	18 (47)	0.5 (*1*)
Female	35 (49)	15 (44)	20 (53)	—[b]
Educational level, n (%)				
No college/university	38 (53)	16 (47)	22 (58)	0.9 (*1*)
College/university	34 (47)	18 (53)	16 (42)	—
Marital status, n (%)				
Married	46 (64)	24 (71)	22 (58)	3.4 (*1*)
Not married	26 (36)	10 (29)	16 (42)	—
Body mass index, mean (SD)	26.4 (4.2)	26.8 (4.2)	26.0 (4.2)	0.8 (*70*)[a]

[a]*t* values with df.

[b]Not applicable.

Intervention Effects on Accelerometer-Based Physical Activity Levels

Results of the multilevel mixed-models repeated measures analyses for accelerometer-based PA are shown in Table 2. A borderline significant intervention effect between baseline and post was found for accelerometer-assessed total PA ($P=.07$). Participants in the intervention group increased their total PA, whereas those in the control group had a decrease in total PA between baseline and post. Similarly, the intervention effect (baseline−follow-up) was borderline significant for accelerometer-based MVPA ($P=.07$); accelerometer-based MVPA increased in the intervention and decreased in the control group.

Table 2. Intervention effects (time-by-group interactions) for objectively assessed physical activity levels (participants with valid accelerometer data in intervention group: baseline=35, post=31, follow-up=32 and control group: baseline=31, post=30, follow-up=27).

Dependent variables (minutes/day)	Baseline (N=66), mean (SD)	Post (N=61), mean (SD)	Follow-up (N=59), mean (SD)	Group×time			
				Reference = Control×pre, beta (SE)	*P* value	Reference = Control×post, beta (SE)	*P* value
Total physical activity (minutes/day)							
Control	283.9 (85.2)	259.8 (71.4)	259.8 (71.4)	Post: 39.4 (19.9)	.05	Follow-up: −24.6 (15.9)	.13
Intervention	273.3 (70.4)	293.7 (85.4)	288.9 (77.5)	Follow-up: 6.5 (6.9)	.35	—[a]	—
Moderate-to-vigorous physical activity (minutes/day)[b]							
Control	29.9 (38.0)	22.1 (14.1)	24.0 (18.3)	Post: 0.8 (0.5)	.13	Follow-up: −0.1 (0.5)	.89
Intervention	17.6 (14.1)	25.6 (30.2)	22.9 (18.9)	Follow-up: 0.4 (0.2)	.07	—	—

[a]Not applicable.

[b]Square-root transformed.

Intervention Effects for Self-Reported Physical Activity Levels

Results of the multilevel mixed-models repeated measures analyses for self-reported domain-specific PA levels are shown in Table 3. For leisure-time PA, (borderline) significant group×time interaction effects were found for vigorous (baseline−follow-up; *P*=.02) and moderate (baseline−post; *P*=.09 and baseline−follow-up; *P*=.01) PA. Leisure-time vigorous PA increased in the intervention and decreased in the control group. For leisure-time moderate PA, the intervention effects were inverse: participants in the control group increased their leisure-time moderate PA, whereas this increased less strongly (baseline−post) or decreased (baseline−follow-up) in the intervention group. For overall leisure-time PA, no significant intervention effects were found. For household-related PA, significant intervention effects were found for moderate PA in the garden (baseline−follow-up; *P*=.04 and post−follow-up; *P*<.001) and moderate household-related PA at home (baseline−post; *P*=.04 and baseline−follow-up; *P*=.01). All intervention effects were in the expected direction: regarding moderate PA in the garden, participants in the intervention group had a steeper increase than participants in the control group. Moderate household-related PA at home increased in the intervention and decreased in the control group. Similarly, a positive intervention effect was found for overall household-related PA (baseline−follow-up; *P*=.05). Finally, a negative intervention effect was found for cycling for transport (post−follow-up; *P*=.07; borderline significant): participants in the control group had a stronger increase in cycling for transport between post and follow-up than participants in the intervention group. For overall transport-related PA, no significant intervention effects were found.

Table 3. Intervention effects (time-by-group interactions) for self-reported domain-specific physical activity (number of participants in the intervention group: baseline=38, post=34, follow-up=34 and control group: baseline=34, post=31, follow-up=31).

Dependent variables (minutes/week)	Baseline (N=72), mean (SD)	Post (N=65), mean (SD)	Follow-up (N=65), mean (SD)	Group×time			
				Reference = Control×pre, beta (SE)	*P* value	Reference = Control×post, beta (SE)	*P* value
Overall leisure-time physical activity							
Control	214.3 (252.3)	324.5 (295.1)	270.2 (272.6)	Post: −3.9 (3.2)	.23	Follow-up: 1.5 (3.3)	.65
Intervention	169.9 (197.2)	220.2 (272.1)	185.9 (198.1)	Follow-up: −2.4 (3.0)	.43	__[a]	—
Leisure-time walking[b]							
Control	156.3 (214.5)	195.8 (226.0)	150.6 (193.1)	Post: −2.3 (2.3)	.32	Follow-up: 2.3 (2.3)	.31
Intervention	92.4 (159.8)	109.6 (209.1)	103.8 (124.9)	Follow-up: −3.8 (9.6)	.97	—	—
Leisure-time vigorous physical activity[b]							
Control	32.6 (146.3)	15.5 (57.9)	15.5 (86.2)	Post: 1.8 (1.1)	.11	Follow-up: 0.4 (1.4)	.77
Intervention	0.00 (0.00)	30.0 (145.0)	22.1 (78.0)	Follow-up: 1.1 (0.5)	.02	—	—
Leisure-time moderate physical activity[b]							
Control	25.3 (55.2)	113.2 (220.1)	104.0 (179.8)	Post: −3.6 (2.1)	.09	Follow-up: −2.1 (2.4)	.39
Intervention	77.5 (153.1)	80.6 (137.0)	60.0 (159.8)	Follow-up: −2.9 (1.0)	.01	—	—
Overall household-related physical activity							
Control	345.6 (292.2)	345.8 (328.5)	385.6 (365.4)	Post: 3.0 (3.3)	.36	Follow-up: 3.6 (3.5)	.30
Intervention	360.8 (360.2)	414.3 (351.2)	603.2 (415.9)	Follow-up: 6.6 (3.4)	.05	—	—
Moderate physical activity garden[b]							
Control	30.8 (60.3)	74.4 (184.6)	82.3 (216.1)	Post: −2.1 (1.9)	.28	Follow-up: 6.7 (1.9)	<.001
Intervention	32.8 (74.3)	17.6 (56.8)	150.0 (212.3)	Follow-up: 2.4 (1.1)	.04	—	—
Vigorous physical activity garden[b]							
Control	17.6 (84.2)	23.2 (129.3)	61.9 (212.1)	Post: 1.5 (2.0)	.44	Follow-up: −3.2 (2.2)	.15
Intervention	37.9 (165.3)	58.2 (168.7)	22.9 (100.2)	Follow-up: −.8 (1.0)	.40	—	—
Moderate physical activity home[b]							
Control	297.1 (290.6)	248.2 (227.4)	241.5 (246.0)	Post: 4.5 (2.2)	.04	Follow-up: 3.1 (2.4)	.19
Intervention	290.1 (309.5)	338.4 (315.6)	430.3 (328.1)	Follow-up: 3.7 (1.4)	.01	—	—
Overall physical activity for transport							
Control	196.6 (228.7)	205.5 (200.8)	277.9 (327.5)	Post: −1.8 (2.8)	.52	Follow-up: −1.7 (2.7)	.52
Intervention	190.8 (284.8)	136.8 (142.4)	153.5 (154.7)	Follow-up: −3.5 (2.9)	.24	—	—
Walking for transport[b]							
Control	148.7 (180.2)	161.6 (178.6)	171.8 (217.8)	Post: −1.0 (2.2)	.66	Follow-up: −0.3 (0.9)	.88
Intervention	133.8 (216.6)	92.9 (122.6)	99.7 (134.7)	Follow-up: −5.9 (1.1)	.59	—	—

Dependent variables (minutes/week)	Baseline (N=72), mean (SD)	Post (N=65), mean (SD)	Follow-up (N=65), mean (SD)	Group×time			
				Reference = Control×pre, beta (SE)	*P* value	Reference = Control×post, beta (SE)	*P* value
Cycling for transport[b]							
Control	47.9 (102.6)	43.9 (84.1)	106.1 (171.7)	Post: .20 (1.3)	.88	Follow-up: −2.3 (1.2)	.07
Intervention	57.0 (137.8)	43.8 (83.2)	53.8 (91.5)	Follow-up: −1.0 (0.8)	.17	—	—

[a]Not applicable.

[b]Square-root transformed.

Discussion

Principal Findings

This study aimed to examine the short- and intermediate-term effects of the *MyPlan 2.0* eHealth intervention on objectively measured and self-reported PA levels in older adults. Regarding objectively measured PA, the results showed that *MyPlan 2.0* had positive but only borderline significant effects for accelerometer-based total PA in the short term and accelerometer-based MVPA in the intermediate term, when support of the website was no longer present. If our findings can be confirmed in a larger study sample, this could suggest that integrating self-regulation principles in behavior change interventions can lead to behavior change [12,33]. The intermediate-term effects found in this study are promising toward health promotion in older adults in the future. Our results are in line with a study by Irvine et al [34], showing intermediate-term effects of an eHealth intervention in adults aged older than 55 years. In that study, most of the positive intervention effects on self-reported PA were maintained after a 3-month period without support from the intervention [34]. To our knowledge, no other study previously examined whether effects of an eHealth intervention on PA in older adults remained after a period without support from the website. The positive intermediate-term effects on MVPA, in the absence of short-term effects, might be explained by the fact that self-regulation can be seen as a goal-guiding process during which individuals are gradually guided toward their goals [12,15]; it might take some time to reach these goals. In addition, it may be that participants start with increasing light-intensity PA, which is more easily achievable and can be reflected in an increase of total PA. When they feel sufficiently comfortable and ready for a *next step*, they might switch to specifically increasing MVPA after a few weeks.

From a health perspective, the effects on accelerometer-based MVPA are very promising. The WHO states that older adults should be physically active for at least 150 min per week [6]. The eHealth intervention *MyPlan 2.0* was able to increase the levels of MVPA in older adults with an average of 5 min per day between baseline and follow-up. This equals an average increase of 35 min per week, which can have a large impact on population health if the intervention would be implemented on a larger scale.

When taking a closer look at the results of objectively assessed PA, it is notable that objectively assessed MVPA decreases between baseline and follow-up in the control group, whereas the intervention group shows an increase. This may be because of seasonal effects. A study by Tucker and Gilliland [35] states that PA levels vary with seasonality. As our baseline measurements took place during an exceptionally warm autumn, post measurements during winter, and follow-up measurements during the beginning of spring (cold and rainy weather), this may be an important reason why MVPA levels decreased between baseline and follow-up in the control group. When linked to the increase in MVPA found in the intervention group, it could be that *MyPlan 2.0* might prevent the seasonal decline in MVPA that is common in older adults, as was observed in the control group. This suggests that, if this study was conducted during 1 season, a greater absolute increase in the intervention group might have been established. Evidently, this is a post hoc explanation and requires further scrutiny.

Besides the effects on objective PA data, this study also investigated the effects of *MyPlan 2.0* on self-reported domain-specific PA. Positive (borderline) significant intervention effects were found for leisure-time vigorous PA, moderate PA in the garden, and moderate household-related PA. However, inverse effects were found for leisure-time moderate PA and cycling for transport. There was no clear consistency in the timing of the positive intervention effects; most effects were found between baseline and follow-up or between post and follow-up, indicating intermediate-term effects. Again, the positive intermediate-term effects, in the absence of short-term effects, might suggest that it takes time for self-regulation techniques to be adopted and used by participants [12,14].

It is important to note that 3 intervention effects were inverse. Cycling for transport (between post and follow-up, borderline significant) and leisure-time moderate PA (both between baseline and post, borderline significant, and between baseline and follow-up, significant) increased more in the control group than the intervention group (small increase or decrease). This suggests that the intervention had a negative effect on these 2 PA domains. It is important to note that users could choose which domain they targeted in their action plans. When examining the content of the action plans in detail, it became clear that no action plans specifically focused on increasing cycling for transport and few action plans (11 out of 59) focused on leisure-time moderate PA (eg, jogging, swimming, and cycling). Of these 11 action plans, 6 focused on both moderate PA and walking. So, the increases in these behaviors in the control group, as opposed to the decreases or less steep increases in the intervention group, may be because of other reasons that

remain unclear until now. A potential other reason could be that some individuals from the control group bought a new bike during the intervention period and, consequently, increased their cycling for transport and/or leisure-time moderate PA. However, this is speculative reasoning as we did not assess whether or not individuals of the control group bought a new bike.

Exploring the content of the action plans in more detail also revealed that many action plans focused on walking during leisure-time (solely or in combination with other leisure-time behaviors: 19/59) and PA at home or in the garden (15/59). This can explain the effects found on household-related PA, but, remarkably, no intervention effects were found on walking during leisure-time. This suggests that participants might not always act upon their proposed action plans. It should be noted that participants were allowed to make more than one action plan, and a previous study showed that participants who formulated multiple action plans focusing on different PA domains were not able to fulfill all these plans [36]. Of course, this is a post hoc reasoning that should be substantiated with data (eg, from personal interviews with participants) to make it possible to draw definite conclusions.

Overall, the finding that effects were mainly found in the intermediate term, when support of the website was no longer present, confirms the practical relevance of developing theory-based eHealth interventions using a self-regulatory perspective. Although our study sample was small, and the results needed to be confirmed in a trial with a longer period without website support, the findings tentatively suggest that eHealth interventions, focusing on pre- and postintentional determinants using specific behavior change techniques, have potential for behavior change in older adults. Consequently, researchers and practitioners should be encouraged to use principles of the self-regulation theory when developing eHealth interventions.

Strengths and Limitations

The main limitation of this study was the low response rate. Of the 1000 retired adults who received a letter, only 5% participated in the study. Additional recruiting (convenience sampling) was needed until the baseline sample of 72 older adults was reached. This indicated that participants were probably very motivated to increase their PA levels, which may have biased the results. Owing to the low response rate, it is not possible to generalize these study results to the general population of older adults. Second, baseline accelerometer-based MVPA differed between the intervention and control group (average difference of 12.3 min per day). Although a multilevel analysis approach was used taking into account clustering of measurements within participants and controlling for baseline PA levels, this large baseline difference in MVPA might have influenced our results. Finally, only retired older adults were included in this study. As PA levels of retired adults might differ from those of working adults, this limits the generalizability of our findings and the comparability with other studies.

The study also had some methodological and theoretical contributions. First of all, the dropout rate was low (10%, 7/72). The study of Degroote et al [20] examining *MyPlan1.0* in adults had a dropout rate of 76% in the intervention group and 56% in the control group. In this study, this was 6.9% in the intervention group and 2.8% in the control group, respectively. This very low dropout might be explained by the improvements that were done to the *MyPlan* website and also by the fact that older adults were targeted in this study. It has been shown that older adults are less likely to drop out from studies than adults [15]. Furthermore, telephone calls were conducted to remind participants to revisit the website when this was not done timely. This might have helped to limit the dropout. However, it is important to keep in mind that follow-up telephone calls might not be feasible when the intervention would be implemented on a larger scale. This may lead to higher attrition rates. Second, objective and self-reported PA levels were measured using validated instruments. Finally, this was one of the first studies to examine the intermediate-term effects of an eHealth intervention when support from the website was no longer available.

Recommendations for Future Research

Future research should examine whether these intervention effects last in the long term in a larger sample. Furthermore, other self-regulation interventions should aim to examine intermediate- and long-term intervention effects instead of focusing mainly on short-term effects. Ideally, the follow-up period should be extended to 6 months to 1 year. In this way, the evidence base on the intermediate- and long-term potential of eHealth interventions aiming to increase PA can be strengthened. To increase the response rate, other recruitment strategies should be used. Recruiting older adults through local service centers, community health centers, and associations for older adults may be more promising than simple random selection through a postal invitation letter. Providing small incentives can also help increase the response rate. Finally, studies using a comparable study protocol and intervention in different countries worldwide should be encouraged. Internet use in older adults differs strongly across countries [37], so it would be useful to discover whether these differences affect the effects and attrition rates of eHealth interventions.

Conclusions

In conclusion, this study adds evidence for the effectiveness of eHealth interventions after a period without support from the website. *MyPlan 2.0* was effective in increasing self-reported leisure-time vigorous PA and moderate household-related PA (home and garden), mainly in the intermediate term when support of the website was no longer present. Although the findings for accelerometer-based MVPA were only borderline significant, this study provided a first indication of the potential of eHealth interventions to increase objectively assessed MVPA in older adults. Future studies with larger samples and long-term follow-up are needed to confirm and clarify these findings.

Acknowledgments

DVD was supported by Research Foundation Flanders (FWO) (grant number: FWO12/PDO/158). The authors would like to thank David Schoukens for his assistance with the data collection and Armand De Clercq for his support in developing *MyPlan 2.0*.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Screenshots of the MyPlan 2.0. website – implementation of self-regulation techniques.
[PDF File (Adobe PDF File), 566 KB - jmir_v21i10e13219_app1.pdf]

Multimedia Appendix 2
CONSORT-eHealth checklist (V1.6.1).
[PDF File (Adobe PDF File), 2282 KB - jmir_v21i10e13219_app2.pdf]

References

1. Bean JF, Vora A, Frontera WR. Benefits of exercise for community-dwelling older adults. Arch Phys Med Rehabil 2004 Jul;85(7 Suppl 3):S31-42; quiz S43. [doi: 10.1016/j.apmr.2004.03.010] [Medline: 15221722]
2. Katzmarzyk PT, Lee I, Martin CK, Blair SN. Epidemiology of physical activity and exercise training in the United States. Prog Cardiovasc Dis 2017;60(1):3-10. [doi: 10.1016/j.pcad.2017.01.004] [Medline: 28089610]
3. McPhee JS, French DP, Jackson D, Nazroo J, Pendleton N, Degens H. Physical activity in older age: perspectives for healthy ageing and frailty. Biogerontology 2016 Jun;17(3):567-580 [FREE Full text] [doi: 10.1007/s10522-016-9641-0] [Medline: 26936444]
4. Musich S, Wang SS, Hawkins K, Greame C. The frequency and health benefits of physical activity for older adults. Popul Health Manag 2017 Jun;20(3):199-207 [FREE Full text] [doi: 10.1089/pop.2016.0071] [Medline: 27623484]
5. Schutzer KA, Graves BS. Barriers and motivations to exercise in older adults. Prev Med 2004 Nov;39(5):1056-1061. [doi: 10.1016/j.ypmed.2004.04.003] [Medline: 15475041]
6. World Health Organization. 2011. Global Recommendations on Physical Activity for Health: 65 Years and Above URL: https://www.who.int/dietphysicalactivity/physical-activity-recommendations-65years.pdf?ua=1 [accessed 2018-12-18] [WebCite Cache ID 74ku7KYTh]
7. Centers for Disease Control and Prevention. 2013. Physical Activity: Data and Statistics URL: https://www.cdc.gov/physicalactivity/data/index.html [accessed 2019-08-19] [WebCite Cache ID 74kuGElGj]
8. Gisle L, Demarest S. Health Interview Survey. 2013. Gezondheidsenquete 2013 - Rapport 2: Gezondheidsgedrag en Leefstijl URL: https://his.wiv-isp.be/nl/Gedeelde%20%20documenten/Summ_LS_NL_2013.pdf [accessed 2018-12-18] [WebCite Cache ID 74kuojHO9]
9. King AC, Rejeski WJ, Buchner DM. Physical activity interventions targeting older adults. A critical review and recommendations. Am J Prev Med 1998 Nov;15(4):316-333. [doi: 10.1016/S0749-3797(98)00085-3] [Medline: 9838975]
10. Cowan LT, van Wagenen SA, Brown BA, Hedin RJ, Seino-Stephan Y, Hall PC, et al. Apps of steel: are exercise apps providing consumers with realistic expectations?: a content analysis of exercise apps for presence of behavior change theory. Health Educ Behav 2013 Apr;40(2):133-139. [doi: 10.1177/1090198112452126] [Medline: 22991048]
11. Lippke S, Ziegelmann JP. Theory-based health behaviour change: developing, testing, and applying theories for evidence-based interventions. Appl Psychol 2008;57(4):698-716. [doi: 10.1111/j.1464-0597.2008.00339.x]
12. Maes S, Karoly P. Self-regulation assessment and intervention in physical health and illness: a review. Appl Psychol 2005 Apr;54(2):267-299. [doi: 10.1111/j.1464-0597.2005.00210.x]
13. Chase JD. Physical activity interventions among older adults: a literature review. Res Theory Nurs Pract 2013;27(1):53-80 [FREE Full text] [doi: 10.1891/1541-6577.27.1.53] [Medline: 23923347]
14. Sniehotta FF, Scholz U, Schwarzer R. Bridging the intention–behaviour gap: planning, self-efficacy, and action control in the adoption and maintenance of physical exercise. Psychol Health 2005 Apr;20(2):143-160. [doi: 10.1080/08870440512331317670]
15. van Dyck D, Plaete J, Cardon G, Crombez G, de Bourdeaudhuij I. Effectiveness of the self-regulation ehealth intervention 'MyPlan1.0' on physical activity levels of recently retired Belgian adults: a randomized controlled trial. Health Educ Res 2016 Oct;31(5):653-664. [doi: 10.1093/her/cyw036] [Medline: 27422898]
16. Plaete J, de Bourdeaudhuij I, Verloigne M, Crombez G. Acceptability, feasibility and effectiveness of an eHealth behaviour intervention using self-regulation: 'MyPlan'. Patient Educ Couns 2015 Jul 26:- (epub ahead of print)(forthcoming)(forthcoming)(forthcoming). [doi: 10.1016/j.pec.2015.07.014] [Medline: 26277282]

17. Ritterband LM, Gonder-Frederick LA, Cox DJ, Clifton AD, West RW, Borowitz SM. Internet interventions: in review, in use, and into the future. Prof Psychol: Res Pract 2003;34(5):527-534. [doi: 10.1037/0735-7028.34.5.527]
18. Collins LM, Murphy SA, Strecher V. The multiphase optimization strategy (MOST) and the sequential multiple assignment randomized trial (SMART): new methods for more potent ehealth interventions. Am J Prev Med 2007 May;32(5 Suppl):S112-S118 [FREE Full text] [doi: 10.1016/j.amepre.2007.01.022] [Medline: 17466815]
19. Cushing CC, Steele RG. A meta-analytic review of ehealth interventions for pediatric health promoting and maintaining behaviors. J Pediatr Psychol 2010 Oct;35(9):937-949. [doi: 10.1093/jpepsy/jsq023] [Medline: 20392790]
20. Degroote L, Plaete J, de Bourdeaudhuij I, Verloigne M, van Stappen V, de Meester A, et al. The effect of the ehealth intervention 'MyPlan 1.0' on physical activity in adults who visit general practice: a quasi-experimental trial. Int J Environ Res Public Health 2018 Jan 30;15(2):E228 [FREE Full text] [doi: 10.3390/ijerph15020228] [Medline: 29385770]
21. Plaete J, de Bourdeaudhuij I, Verloigne M, Oenema A, Crombez G. A self-regulation ehealth intervention to increase healthy behavior through general practice: protocol and systematic development. JMIR Res Protoc 2015 Dec 22;4(4):e141 [FREE Full text] [doi: 10.2196/resprot.4835] [Medline: 26694101]
22. Muellmann S, Forberger S, Möllers T, Bröring E, Zeeb H, Pischke CR. Effectiveness of ehealth interventions for the promotion of physical activity in older adults: a systematic review. Prev Med 2018 Mar;108:93-110. [doi: 10.1016/j.ypmed.2017.12.026] [Medline: 29289643]
23. Kannus P. Preventing osteoporosis, falls, and fractures among elderly people. Promotion of lifelong physical activity is essential. Br Med J 1999 Jan 23;318(7178):205-206 [FREE Full text] [doi: 10.1136/bmj.318.7178.205] [Medline: 9915707]
24. Rzewnicki R, Auweele YV, de Bourdeaudhuij I. Addressing overreporting on the international physical activity questionnaire (IPAQ) telephone survey with a population sample. Public Health Nutr 2003 May;6(3):299-305. [doi: 10.1079/PHN2002427] [Medline: 12740079]
25. Poppe L, van der Mispel C, de Bourdeaudhuij I, Verloigne M, Shadid S, Crombez G. Users' thoughts and opinions about a self-regulation-based ehealth intervention targeting physical activity and the intake of fruit and vegetables: a qualitative study. PLoS One 2017;12(12):e0190020 [FREE Full text] [doi: 10.1371/journal.pone.0190020] [Medline: 29267396]
26. Vandelanotte C, Spathonis KM, Eakin EG, Owen N. Website-delivered physical activity interventions a review of the literature. Am J Prev Med 2007 Jul;33(1):54-64. [doi: 10.1016/j.amepre.2007.02.041] [Medline: 17572313]
27. Craig CL, Marshall AL, Sjöström M, Bauman AE, Booth ML, Ainsworth BE, et al. International physical activity questionnaire: 12-country reliability and validity. Med Sci Sports Exerc 2003 Aug;35(8):1381-1395. [doi: 10.1249/01.MSS.0000078924.61453.FB] [Medline: 12900694]
28. van Dyck D, Cardon G, Deforche B, de Bourdeaudhuij I. IPAQ interview version: convergent validity with accelerometers and comparison of physical activity and sedentary time levels with the self-administered version. J Sports Med Phys Fitness 2015;55(7-8):776-786. [Medline: 24921615]
29. International Physical Activity Questionnaires (IPAQ). 2015. Guidelines for Data Processing and Analysis of the International Physical Activity Questionnaire URL: http://www.ipaq.ki.se [accessed 2019-08-28]
30. Freedson PS, Melanson E, Sirard J. Calibration of the Computer Science and Applications Inc accelerometer. Med Sci Sports Exerc 1998 May;30(5):777-781. [doi: 10.1097/00005768-199805000-00021] [Medline: 9588623]
31. Bates D, Maechler M, Bolker B, Walker S. The Comprehensive R Archive Network. 2014. lme4: Linear Mixed-Effects Models using 'Eigen' and S4 URL: http://CRAN.R-project.org/package=lme4 [accessed 2018-12-20] [WebCite Cache ID 74kvQPHVj]
32. Chakraborty H, Gu H. RTI International. 2009. A Mixed Model Approach for Intent-to-Treat Analysis in Longitudinal Clinical Trials With Missing Values URL: https://www.rti.org/rti-press-publication/mixed-model-approach-intent-treat [accessed 2018-12-20]
33. Cleary TJ, Zimmerman BJ. Self-regulation empowerment program: a school-based program to enhance self-regulated and self-motivated cycles of student learning. Psychol Sch 2004 May;41(5):537-550. [doi: 10.1002/pits.10177]
34. Irvine AB, Gelatt VA, Seeley JR, Macfarlane P, Gau JM. Web-based intervention to promote physical activity by sedentary older adults: randomized controlled trial. J Med Internet Res 2013 Feb 5;15(2):e19 [FREE Full text] [doi: 10.2196/jmir.2158] [Medline: 23470322]
35. Tucker P, Gilliland J. The effect of season and weather on physical activity: a systematic review. Public Health 2007 Dec;121(12):909-922. [doi: 10.1016/j.puhe.2007.04.009] [Medline: 17920646]
36. Plaete J, de Bourdeaudhuij I, Verloigne M, Crombez G. The use and evaluation of self-regulation techniques can predict health goal attainment in adults: an explorative study. PeerJ 2016;4:e1666 [FREE Full text] [doi: 10.7717/peerj.1666] [Medline: 26966648]
37. European Commission. 2015. People in the EU: Who Are We and How Do We Live? URL: https://ec.europa.eu/eurostat/en/web/products-statistical-books/-/KS-04-15-567 [accessed 2018-12-20] [WebCite Cache ID 78MfRG2PS]

Abbreviations

cpm: counts per minute
eHealth: electronic health

FWO: Research Foundation Flanders
IPAQ: International Physical Activity Questionnaire
MVPA: moderate-to-vigorous physical activity
PA: physical activity
WHO: World Health Organization

Edited by CL Parra-Calderón; submitted 21.12.18; peer-reviewed by A Maier, B Xie, K Knittle, J Boekhout-Oppeneer; comments to author 31.03.19; revised version received 15.05.19; accepted 09.07.19; published 07.10.19.

Please cite as:
Van Dyck D, Herman K, Poppe L, Crombez G, De Bourdeaudhuij I, Gheysen F
Results of MyPlan 2.0 on Physical Activity in Older Belgian Adults: Randomized Controlled Trial
J Med Internet Res 2019;21(10):e13219
URL: https://www.jmir.org/2019/10/e13219
doi:10.2196/13219
PMID:31593541

Original Paper

Cost and Effectiveness of Blended Versus Standard Cognitive Behavioral Therapy for Outpatients With Depression in Routine Specialized Mental Health Care: Pilot Randomized Controlled Trial

Lisa Catharine Kooistra[1,2,3], MSc; Jenneke Elize Wiersma[2,3], PhD; Jeroen Ruwaard[2,3], PhD; Koen Neijenhuijs[1,3,4], MSc; Joran Lokkerbol[5], PhD; Patricia van Oppen[2,3,6], PhD; Filip Smit[1,3,5,7], PhD; Heleen Riper[1,2,3,6], PhD

[1]Department of Clinical, Neuro and Developmental Psychology, Clinical Psychology Section, Vrije Universiteit Amsterdam, Amsterdam, Netherlands
[2]Department of Research and Innovation, GGZ in Geest/Amsterdam University Medical Center, VU University Medical Center, Amsterdam, Netherlands
[3]Amsterdam Public Health Research Institute, Amsterdam University Medical Center, VU University Medical Center, Amsterdam, Netherlands
[4]Cancer Center Amsterdam, Amsterdam University Medical Center, VU University Medical Center, Amsterdam, Netherlands
[5]Center of Economic Evaluation, Trimbos Institute (Netherlands Institute of Mental Health and Addiction), Utrecht, Netherlands
[6]Department of Psychiatry, GGZ in Geest/Amsterdam University Medical Center, VU University Medical Center, Amsterdam, Netherlands
[7]Department of Epidemiology and Biostatistics, Amsterdam University Medical Center, VU University Medical Center, Amsterdam, Netherlands

Corresponding Author:
Lisa Catharine Kooistra, MSc
Department of Research and Innovation
GGZ in Geest/Amsterdam University Medical Center, VU University Medical Center
Postbus 74077
Amsterdam, 1070 BB
Netherlands
Phone: 31 204448180
Email: l.kooistra@ggzingeest.nl

Abstract

Background: Cognitive behavioral therapy (CBT) is an effective treatment, but access is often restricted due to costs and limited availability of trained therapists. Blending online and face-to-face CBT for depression might improve cost-effectiveness and treatment availability.

Objective: This pilot study aimed to examine the costs and effectiveness of blended CBT compared with standard CBT for depressed patients in specialized mental health care to guide further research and development of blended CBT.

Methods: Patients were randomly allocated to blended CBT (n=53) or standard CBT (n=49). Blended CBT consisted of 10 weekly face-to-face sessions and 9 Web-based sessions. Standard CBT consisted of 15 to 20 weekly face-to-face sessions. At baseline and 10, 20, and 30 weeks after start of treatment, self-assessed depression severity, quality-adjusted life-years (QALYs), and costs were measured. Clinicians, blinded to treatment allocation, assessed psychopathology at all time points. Data were analyzed using linear mixed models. Uncertainty intervals around cost and effect estimates were estimated with 5000 Monte Carlo simulations.

Results: Blended CBT treatment duration was mean 19.0 (SD 12.6) weeks versus mean 33.2 (SD 23.0) weeks in standard CBT (P<.001). No significant differences were found between groups for depressive episodes (risk difference [RD] 0.06, 95% CI −0.05 to 0.19), response to treatment (RD 0.03, 95% CI −0.10 to 0.15), and QALYs (mean difference 0.01, 95% CI −0.03 to 0.04). Mean societal costs for blended CBT were €1183 higher than standard CBT. This difference was not significant (95% CI −399 to 2765). Blended CBT had a probability of being cost-effective compared with standard CBT of 0.02 per extra QALY and 0.37 for an additional treatment response, at a ceiling ratio of €25,000. For health care providers, mean costs for blended CBT were €176 lower than standard CBT. This difference was not significant (95% CI −659 to 343). At €0 per additional unit of effect, the probability of blended CBT being cost-effective compared with standard CBT was 0.75. The probability increased to 0.88 at a ceiling ratio of €5000 for an added treatment response, and to 0.85 at €10,000 per QALY gained. For avoiding new depressive episodes, blended CBT was deemed not cost-effective compared with standard CBT because the increase in costs was associated with negative effects.

Conclusions: This pilot study shows that blended CBT might be a promising way to engage depressed patients in specialized mental health care. Compared with standard CBT, blended CBT was not considered cost-effective from a societal perspective but had an acceptable probability of being cost-effective from the health care provider perspective. Results should be carefully interpreted due to the small sample size. Further research in larger replication studies focused on optimizing the clinical effects of blended CBT and its budget impact is warranted.

Trial Registration: Netherlands Trial Register NTR4650; https://www.trialregister.nl/trial/4408

International Registered Report Identifier (IRRID): RR2-10.1186/s12888-014-0290-z

(J Med Internet Res 2019;21(10):e14261) doi:10.2196/14261

KEYWORDS

depression; blended cognitive behavioral therapy; specialized mental health care; cost-effectiveness; randomized controlled trial

Introduction

Several evidence-based pharmacological and psychological treatments have been developed for major depressive disorder (MDD) [1,2]. Within the domain of psychotherapy, there is an especially large body of evidence supporting the efficacy of cognitive behavioral therapy (CBT) [3]. Unfortunately, there is a discrepancy between treatment availability and treatment demand [4]. This can partly be explained by increasingly insufficient mental health care budgets, which limit the availability of trained psychotherapists who can provide evidence-based treatments [5]. As a consequence, patients with severe symptoms often do not receive treatment, or they have to be placed on waiting lists rather than having immediate access to specialized depression care [4,6]. Therefore, there is a high need for efficient and cost-effective mental health care to manage this problem [7].

Web-based (online) psychotherapy is often cited as a promising way to reduce treatment costs and increase treatment availability for common mental disorders [8-11]. For example, it may lower the required therapist time per patient [12]. The efficacy of several online treatments has been demonstrated for the treatment of depression when compared with control groups [13-15]. Evidence suggests that therapist-guided online treatment can be equally as effective as standard face-to-face therapy [16,17]. The cost-effectiveness of guided online treatment has been less well studied. An individual patient data meta-analysis, combining data from five randomized controlled trials (RCTs), suggested that guided online treatment is not yet cost-effective compared with control conditions [18].

To date, most studies have focused on community samples and self-referred participants. Less is known about the costs and effects of online therapy in routine specialized mental health care for patients with more severe or complex depression profiles. This patient group is likely to require more personalized treatment, such as monitoring of suicidal ideation or addressing comorbid disorders [16]. The integration of online and face-to-face treatment into a blended treatment format could be a promising way to address these matters [19-23]. Blended treatment allows therapists to closely monitor their patients, both in face-to-face sessions at the clinic and in an online environment. At the same time, it is thought to retain the positive aspects associated with online treatment, such as lower costs compared with standard treatment, reduction of therapist time, and increased patient self-management [19,24,25].

Although several studies are in progress, current evidence for blended treatment is limited [26-29]. Most studies to date have evaluated the online component as an add-on to standard care rather than offering an integrated blended treatment protocol [30-35]. Results of these studies are promising and indicate that blended treatment may be a viable treatment option. However, the treatment format limits the margin for cost-effectiveness because therapist time is not reduced and overall treatment dosage may even increase [18,36]. Recently, Thase and colleagues [25] used a more integrated approach to blended treatment. Therapist time was limited by combining nine online sessions with twelve 25-minute (instead of 50-minute) face-to-face sessions. Compared with CBT (n=77), blended CBT (n=77) led to noninferior results on the Hamilton Depression Rating Scale (HAM-D) in medication-free adults with depression, when provided in two university clinics.

This study compares the costs and effects of integrated blended CBT with standard CBT when provided to depressed patients in the acute phase of treatment in specialized mental health care. The blended treatment aimed to replace half of face-to-face treatments with online sessions, and thereby shorten treatment duration when compared with standard CBT for MDD [22]. The study examines whether blended CBT has the potential to lead to comparable clinical effects as standard CBT at lower costs. The study was designed as a pilot study with the aim to guide further development of blended CBT and inform future research on feasibility, effectiveness, and cost-effectiveness of blended CBT for depression in outpatient specialized mental health care.

Methods

Study Design and Participants

This pilot study was designed as an RCT in three mental health care organizations in the Netherlands, at five outpatient treatment locations. Blended CBT was compared with standard CBT for depression. Outcomes were measured before treatment allocation at baseline and 10, 20, and 30 weeks after start of treatment. Adult patients with a *Diagnostic and Statistical Manual of Mental Disorders* (Fourth Edition, Text Revision; *DSM-IV-TR*) diagnosis [37] of MDD, who were indicated for individual CBT by the local intake staff, were recruited during the intake procedure. In the Netherlands, patients can be referred

to specialized services when they do not respond to treatment in primary mental health care [38]. Therefore, patients in this trial were likely to have complex or severe clinical profiles and to have undergone some form of pharmacotherapy or short-term psychotherapy before this study.

Exclusion criteria were inadequate proficiency in the Dutch language; no valid email address or no computer with internet access; current psychotic disorder, bipolar disorder, or substance dependence; or high risk for suicide (current plans). Patients with current substance dependence could participate in the trial if they reported abstinence and had been treated at a center specialized in treatment of substance abuse for at least a month before random allocation.

Diagnoses and suicidal ideation were assessed with the Mini-International Neuropsychiatric Interview Plus (MINI-Plus) [39,40]. Assessors were trained research assistants with master's degrees in psychology. The study protocol has been published [41]. The trial was registered in the Netherlands Trial Register (registration number NTR4650) and approved by the Medical Ethics Committee of the VU University Medical Center Amsterdam (registration number 2014.191).

Randomization and Masking

Patients were randomly allocated to blended CBT or standard CBT when they met the inclusion criteria, provided a signed informed consent form, and completed the full baseline assessment. Allocation was stratified by mental health care center using a computer-generated random number table. An independent researcher conducted the randomization. Patients and therapists were aware of group allocation, but allocation was concealed from outcome assessors.

Interventions

Both interventions consisted of cognitive behavioral depression treatment. The content of CBT was based on the standard treatment manual [42-44], which recommends 15 to 20 weekly sessions focusing on psychoeducation, behavioral activation, cognitive restructuring, and relapse prevention. As part of the treatment, patients in both treatment groups were encouraged to fill in the 16-item Quick Inventory of Depressive Symptomatology Self-Report on a weekly basis to allow patients and their therapists to monitor change in depression severity [45]. Patients who completed 14 sessions or more (75% of 18 sessions) were considered treatment completers. Parallel pharmacological treatment was allowed.

Blended CBT consisted of 10 face-to-face sessions at the specialized mental health care center and 9 Web-based (online) sessions that patients worked through at home. Therapists and patients had personal, password-protected accounts on the Web-based treatment platform. Therapists were trained in the use of the Web-based treatment platform and received a treatment manual. Patients received information on how to work with the platform in the first face-to-face session. After each online session, therapists provided online therapeutic feedback regarding content and progress. Therapists could let patients repeat an online session if this was warranted. The blended CBT protocol aimed to provide one face-to-face session, one online session, and one online feedback message per week over 10 weeks, starting with a face-to-face session. The intervention was semistructured, offering online sessions in a fixed order and providing therapists with a protocol for each face-to-face session. Therapists were allowed to personalize face-to-face sessions to each patient's individual needs and situation in terms of themes and techniques. More detailed information on the format and content of blended CBT can be found elsewhere [22,41].

The standard CBT group received individual face-to-face cognitive behavioral depression treatment at the clinic in accordance with routine care procedures. The content of standard CBT was comparable to the content of blended CBT. Therapists were advised to plan weekly sessions but were allowed to deviate from the treatment manual when necessary. Based on the guideline, standard CBT was expected to consist of 18 sessions on average, provided over approximately 20 weeks. However, duration of treatment could vary per patient.

Outcomes

The primary clinical outcome was self-reported depression severity at 10, 20, and 30 weeks after the start of treatment, as measured by a Web-based version of the Inventory of Depressive Symptomatology Self-Report ($IDS\text{-}SR_{30}$) [46-49]. Within the cost-effectiveness framework, the $IDS\text{-}SR_{30}$ scores were used to assess treatment response based on the reliable change index (RCI) [50]. Baseline $IDS\text{-}SR_{30}$ scores were subtracted from follow-up scores and then divided by the standard error of the difference scores (SE 4.78, Cronbach alpha=.84). Treatment response (reliable change) was coded as absent ($RCI \geq -1.96$) versus present ($RCI < -1.96$). Remission was defined as the combination of an RCI less than −1.96 and an $IDS\text{-}SR_{30}$ score less than 13, indicating no depression severity on the $IDS\text{-}SR_{30}$ severity index [46,51].

Clinician-rated outcomes included the presence of a depressive episode at 10, 20, and 30 weeks after the start of treatment and during follow-up. Depression diagnosis was assessed at 10, 20, and 30 weeks after the start of treatment during a telephone interview with section A of the MINI-Plus diagnostic interview [39,40] by trained research assistants who were blinded to treatment allocation. At baseline and 30-week follow-up, all sections of the MINI-Plus interview were administered during a face-to-face or telephone interview to assess comorbid diagnoses.

To calculate quality-adjusted life-years (QALYs), the EQ-5D three-level version [52] questionnaire was administered online at all four time points. QALYs gained over the 30-week study period were estimated by linearly interpolating EQ-5D utility over the time points and correcting for time. Utility scores were based on the Dutch tariff [53].

Self-reported resource use in the 4 weeks before each assessment was measured with a Web-based version of the Trimbos/iMTA questionnaire for Costs associated with Psychiatric Illness (TiC-P) [54]. The TiC-P includes questions on (1) the number of visits to various health care providers in primary and specialized care settings; (2) medication use for sleep, depression, and anxiety; (3) help from friends and family; and (4) absenteeism and reduced work productivity (presenteeism)

in paid and unpaid work. For the main economic evaluation, costs were estimated from a societal perspective, which included all costs measured by the TiC-P. The health care provider perspective only included direct medical costs. A full overview of all cost categories can be found in Multimedia Appendix 1. Costs were computed by multiplying the number of units (contacts, visits, sessions) by the standard unit cost prices (for the year 2014) as reported in the most recent Dutch guideline for economic evaluations [55]. When costs were not included in the latest version of the manual, the cost price was derived from the previous guideline [56] and indexed to the year 2014. Medication prices were retrieved from the National Healthcare Institute (Zorginstituut Nederland) (Z-index [57]). For each prescription, the standard dispensing fee for pharmacists of €6 was added. For blended CBT patients, costs associated with the actual number of online feedback messages received were added, based on an estimated 30 minutes of therapist time per feedback message. Cumulative costs over the 30-week study period were estimated using linear interpolation.

Statistical Analyses

For descriptive purposes, clinical and cost outcomes were first examined in separate linear mixed-effect models. To account for missing data and the correlation between follow-up time points, linear mixed-effect models with restricted maximum likelihood were used to estimate the treatment effects, EQ-5D utility gained, and cumulative costs across time. The models included both fixed and random effects, which allowed for estimation on a patient level of how far each patient diverged from the fixed (group level) effect over time [58,59]. Main outcome measures (ie, IDS-SR depression severity, reliable change, MINI depression status, EQ-5D QALYs, and costs) were evaluated in separate mixed models. Time was included as a categorical variable (0, 10, 20, and 30 weeks). A logistic mixed-effects model was fitted for diagnosis of depressive episodes (present versus absent) and reliable improvement of depression severity (improvement versus no improvement). In these models, the random intercept was dropped from the model because all patients started at the same baseline value (a current diagnosis of depression).

For the cost-effectiveness analyses, baseline adjusted linear mixed-effect models were estimated with group as the independent variable and societal costs, direct medical costs, and reliable change in depression severity as the dependent variables and a random effect across individuals. For dichotomous outcomes, linear mixed-effect models were fitted to estimate risk difference and calculate the incremental cost-effectiveness ratio (ICER). EQ-5D QALYs gained during the study were used as the dependent variable for the cost-utility analysis. Within the cost-effectiveness and cost-utility framework, time was not included in the regression models, rendering one estimate for cost, effectiveness, or utility for all follow-ups. Societal costs during the study period were used as the main cost outcome. Direct health care costs were examined as a separate cost outcome to conduct the health-economic evaluation from the health care system perspective. Uncertainty intervals of 95% around linear mixed-effect model regression estimates for group were estimated with 5000 probabilistic Monte Carlo simulations, running all mixed models simultaneously.

The resulting pairs of cost and effect or utility estimates were plotted on cost-effectiveness and cost-utility planes. The y-axis represented relative costs, and the x-axis represented relative effects associated with blended CBT versus standard CBT. The axes divide the plane into four quadrants. In the northeast quadrant, blended CBT was more expensive and more effective than standard CBT. In the northwest quadrant, blended CBT was more expensive and less effective than status quo (standard CBT), meaning that blended CBT was dominated by standard CBT. In the southwest quadrant, blended CBT was less expensive and less effective than standard CBT. In the southeast quadrant, blended CBT was less expensive and more effective than standard CBT, which indicated that blended CBT dominated standard CBT. Cost-effectiveness acceptability curves were estimated to assess the probability of blended CBT being cost-effective given various willingness-to-pay ceilings.

The mixed models were estimated using the lmer and glmer functions from the lme4 package (version 1.1-15) [60] using R software (version 3.4.4) [61]. Combined cost-effectiveness and cost-utility analyses were performed in Stata version 14.2 (StataCorp LP, College Station, TX, USA).

Two sensitivity analyses were conducted. First, a linear mixed-effect model was estimated for societal costs excluding the costs of in-patient care (both in general hospitals and mental health care). Although these events are rare, they can act as influential outliers due to the high costs associated with in-patient care. Second, direct nonmedical costs associated with absenteeism and presenteeism were evaluated as a separate cost outcome to assess their contribution to the overall societal costs as evaluated in the main analysis.

Power

Sample size estimation was based on the probability of blended CBT being cost-effective in comparison with standard CBT for various willingness-to-pay ceilings (ie, the maximum additional financial contribution society is willing to invest to gain one more unit of treatment effect) [62]. Through a simulation study, the impact of different sample sizes on the stability of the cost-effectiveness acceptability curve was determined [62]. Using realistic fixed estimates of the mean and standard deviation of effects and the costs on the population level (based on Hakkaart-van Roijen et al [63]), a large number of trials were simulated in which sample sizes were systematically varied between n=10 to n=500 per group. At a sample size of n=75 per group, the probability estimates converged to acceptable 75% values within the relevant range of willingness-to-pay ceilings. Therefore, this study aimed for a sample size of N=150.

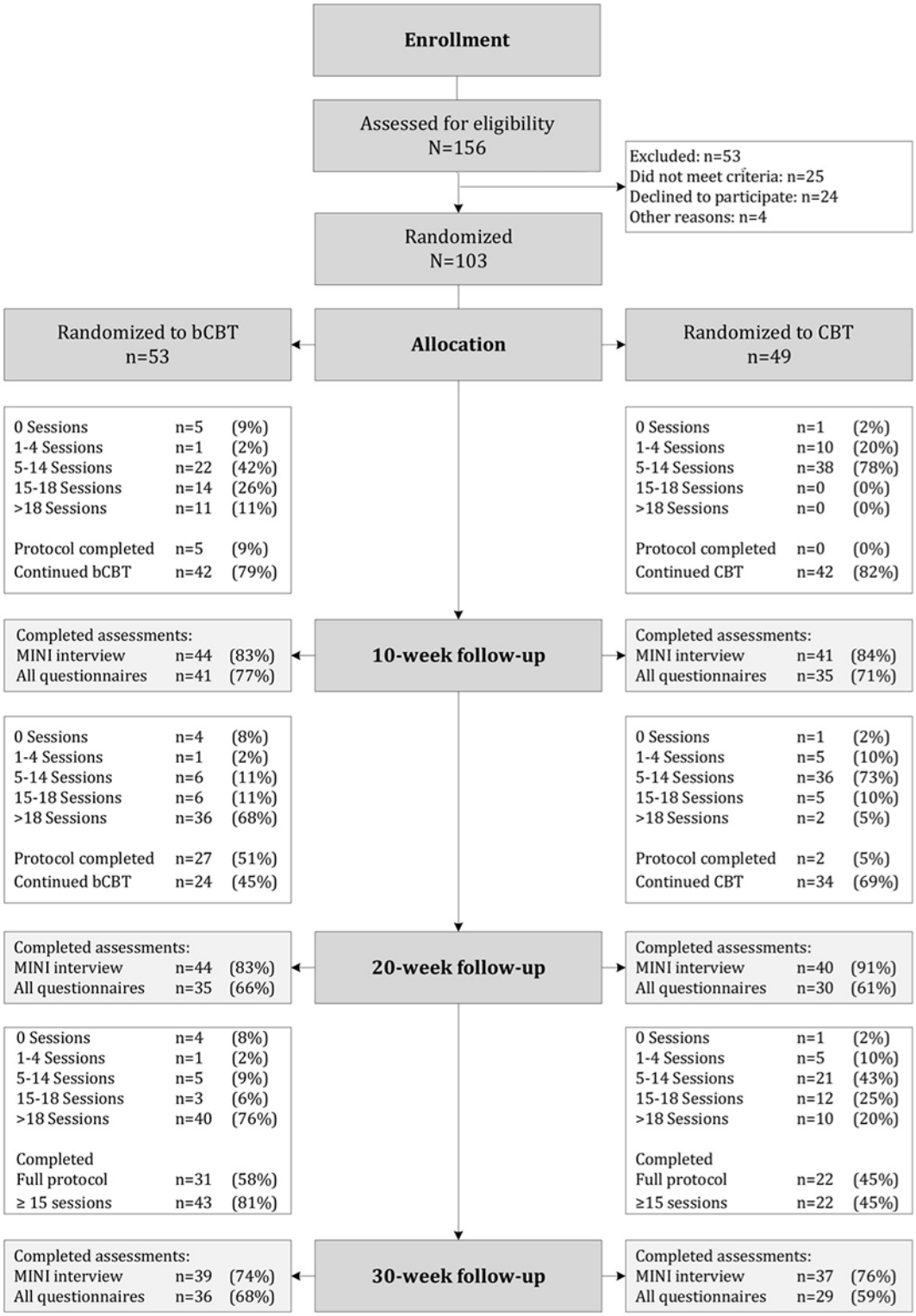
Results

Overview

Between August 2014 and May 2016, 103 patients were randomized to blended CBT (n=54) or standard CBT (n=49). One patient withdrew from the study before commencing treatment due to starting another treatment elsewhere. The trial profile is presented in Figure 1. The full intention-to-treat sample

consisted of 102 patients (blended CBT: n=53; standard CBT: n=49). This sample size was smaller than the initial goal of 150 patients [41]. The recruitment and screening of sufficient patients to randomize 150 patients proved to be unfeasible within the fixed time frame of this study. Fewer patients were referred to specialized services than expected. This was partly due to a nationwide reorganization of Dutch mental health care, which meant that a larger proportion of depressed patients were treated in primary mental health care rather than specialized care.

Figure 1. Flowchart. bCBT: blended cognitive behavioral therapy; CBT: standard cognitive behavioral therapy.



Information on patients' baseline demographic characteristics, clinical profiles, and costs are presented in Table 1. A notable number of patients (69 of 102, 68%) were diagnosed with at least one other *DSM-IV* disorder in addition to MDD, 60% (61 of 102) reported suicidal ideation, and 28% (29 of 102) had attempted suicide at some point in their lives. Patients in blended CBT showed a higher baseline depression severity (45.2 versus 41.5), lower mean utility score (0.36 versus 0.46), and higher

direct medical costs (€989 versus €439) in the 4 weeks before baseline than patients in standard CBT.

Table 1. Sample characteristics of participants at baseline (N=102).

Patient characteristics	Blended CBT[a] (n=53)	Standard CBT (n=49)	Total (N=102)
Demographic			
Gender (female), n (%)	35 (66)	29 (60)	64 (63)
Age (years), mean (SD)	39.3 (11.3)	38.1 (10.6)	38.8 (10.9)
In a relationship, n (%)	30 (57)	30 (61)	60 (59)
Education, n (%)			
Low	6 (11)	2 (4)	8 (8)
Middle	30 (57)	33 (67)	63 (62)
High	17 (32)	14 (29)	31 (30)
Employed, n (%)	31 (59)	28 (57)	59 (58)
Nationality (Dutch), n (%)	48 (91)	47 (96)	95 (93)
Clinical characteristics			
Blended CBT treatment preference, n (%)	33 (62)	33 (67)	66 (65)
Any comorbidity,[b] n (%)	38 (72)	31 (63)	69 (68)
Anxiety disorders, n (%)	30 (57)	25 (51)	55 (54)
Other disorders,[c] n (%)	25 (47)	23 (47)	48 (47)
Comorbid disorders,[d] mean (SD)	1.8 (1.7)	1.8 (1.9)	1.8 (1.7)
MDD[e] duration in months, mean (SD)	73.7 (73.2)	80.2 (69.8)	76.9 (70.9)
Depression (IDS-SR)[f], mean (SD)	45.2 (12.1)	41.5 (11.6)	43.4 (11.9)
Severity (IDS-SR), n (%)			
Mild/moderate	17 (32)	22 (45)	39 (38)
Severe/very severe	36 (68)	27 (55)	64 (63)
Utility scores, mean (SD)	0.36 (0.29)	0.46 (0.31)	0.41 (0.30)
Prior treatment (4 weeks), n (%)			
Antidepressant medication	36 (68)	29 (59)	65 (64)
Psychological treatment	44 (83)	42 (86)	86 (84)
Costs (€, TiC-P[g], 4 weeks), mean (SD)			
Direct medical	989 (1626)	439 (398)	725 (1230)
Direct nonmedical	428 (802)	246 (526)	339 (685)
Indirect nonmedical	796 (1629)	716 (1610)	757 (1613)
Societal costs	2205 (2615)	1401 (1762)	1819 (2271)

[a]CBT: cognitive behavioral therapy.

[b]The *DSM-IV-TR* and *ICD-10* codes are available on request.

[c]Obsessive compulsive disorder, posttraumatic stress disorder, alcohol/drug abuse, somatoform disorders.

[d]Anxiety disorders: social phobia, panic with or without agoraphobia, agoraphobia, generalized anxiety disorder.

[e]MDD: major depressive disorder.

[f]IDS-SR: Inventory of Depressive Symptomatology, self-report version.

[g]TiC-P: Trimbos/iMTA questionnaire for Costs associated with Psychiatric Illness.

Study Dropout

Overall, the MINI-Plus diagnostic interview could be administered to 85 of 102 patients (83%) at 10 weeks, 84 patients (82%) at 20 weeks, and 76 patients (75%) at 30 weeks. The online self-report questionnaires were completed by 77 patients at 10 weeks (76%), 65 patients (64%) at 20 weeks, and 65 patients (64%) at 30 weeks.

Patients with missing data at one or more follow-up assessments (n=49) were on average five years younger than patients who completed all assessments (n=54; mean 36.0, SD 10.5 years versus mean 41.2, SD 10.8 years; t_{100}=2.49, *P*=.02). Within the blended CBT group, missing data were not associated with patient characteristics. Patients with missing data in the standard CBT group were more often unemployed than patients who did not have missing data (χ^2_2=14.1, *P*<.001). Figure 1 displays study and treatment adherence for both groups throughout the study period.

Treatment Adherence

Of 102 patients, 97 (95%) started treatment and received at least one session. In blended CBT, the average number of sessions matched the planned number of 19.5 sessions (SD 8.3, range 0 to 31). On average, blended CBT had a mean of 9.6 online sessions (SD 4.4, range 0 to 16) and 10.0 face-to-face sessions (SD 4.6, range 0 to 16). Per-patient online feedback was provided a mean of 8.4 times (SD 4.2, range 0 to 15). Based on the cutoff of 14 sessions (75% of the optimal blended CBT protocol of 19 sessions), 43 of 53 blended CBT patients (81%) were considered treatment completers. When the 75% completion criterion was applied to online and face-to-face sessions separately (>7 each), 40 blended CBT patients (75%) were considered completers. Seven patients (13%) completed the full blended CBT protocol within 10 weeks.

In standard CBT, the mean number of sessions was 13.3 (SD 6.3, range 0 to 27), which was less than planned (15-18 sessions) within the time frame of this study. Based on the cutoff of 14 sessions, 22 of 49 patients (45%) were considered treatment completers. Four patients (8%) completed the standard CBT protocol (16-20 sessions) within 20 weeks.

Compared with patients in the standard CBT group, blended CBT patients received significantly more sessions in total (19 versus 13, t_{100}=−4.09, *P*<.001), but significantly fewer face-to-face sessions at the clinic (10 versus 13, t_{100}=3.07, *P*=.003). As expected based on the treatment protocol, mean treatment duration was significantly shorter in blended CBT than in standard CBT, with an average of 19.0 (SD 12.6) weeks versus 33.2 (SD 23.0) weeks (t_{100}=3.91, *P*<.001). When therapist time spent on online feedback was included, the combined amount of direct and indirect therapist time did not differ between groups with an average of 14.0 (SD 6.2) hours for blended CBT and 13.3 (SD 6.3) hours for standard CBT (t_{100}=−0.55, *P*=.58).

Clinical Outcomes

Deviation between the desired time points and actual time of data collection was deemed within an acceptable range to enter time as a fixed categorical variable (baseline or 0 weeks, and 10, 20, and 30 weeks after start of treatment) in the linear mixed models (see Multimedia Appendix 2). No differences were found between groups in psychotropic medication use at all assessment periods. Controlling for demographic variables associated with missing data (age and employment status) did not improve model fit in all models (*P*>.05). Controlling for baseline scores did significantly improve model fit in all models with continuous outcomes (*P*<.001). Unadjusted (observed) means and outcomes per follow-up assessment are presented in Table 2. Controlling for baseline severity, no group difference was found in decrease of depression severity over time on the IDS-SR_{30} (overall: *b*=2.03, 95% CI −1.57 to 5.64; $t_{50.45}$=1.17, *P*=.25). Time significantly predicted a decrease of depression severity at all time points (overall: *b*=−7.13, 95% CI −9.64 to −4.71; $t_{49.55}$=−5.74, *P*<.001). For the full sample, the estimated mean depression severity decreased from severe (43.23, 95% CI 42.07-44.40) to moderate (27.1, 95% CI 22.03-32.14) [45,48].

Controlling for baseline utility scores, time significantly predicted an increase in QALYs gained at all time points (overall, *b*=.12, 95% CI 0.10-0.13; $t_{73.84}$= 13.77, *P*<.001). There was no significant difference in QALY gains between groups (*b*=−0.01, 95% CI −0.03 to 0.01; $t_{73.74}$=−0.88, *P*=.38).

For reliable change in depression severity (treatment response), a significant association was found between time and treatment response (*b*=1.81, OR 6.09, 95% CI 0.89-2.89; *z*=4.00, *P*<.001), but no significant difference between groups (*b*=−0.31, OR 0.73, 95% CI −1.61 to 0.99; *z* =−0.53, *P*=.60). Twenty patients reported significant improvement at all follow-up assessments (blended CBT: n=10; standard CBT: n=10). There were no patients who consecutively reported deterioration at all follow-up assessments. No deaths occurred during the study. One depression-related, but not treatment-related, serious adverse event occurred in the standard CBT group, in which one person self-harmed. Table 3 provides an overview of reliable change compared with baseline per follow-up, per group.

The proportion of patients who did not fulfill criteria for a depressive episode at 10, 20, and 30 weeks after the start of treatment did not differ significantly between groups (overall: *b*=.64, OR 1.89, 95% CI 0.73-4.90; *z*=1.31, *P*=.19). Odds of a depressive episode decreased significantly over time (overall: *b*=−1.44, OR 0.23, 95% CI 0.16-0.35; *z*=−7.12, *P*<.001). At 30-week follow-up, 23 of 102 patients (31% of full sample) met criteria for remission, reporting an absence of symptoms for at least 8 weeks. No significant between-group differences were found in the number of comorbid diagnoses. Table 4 provides an overview of the current diagnoses at 30-week follow-up for both groups.

Table 2. Unadjusted (observed) means and primary clinical and utility outcomes.

Outcome	Blended CBT[a]		Standard CBT		Full sample		Blended CBT vs standard CBT
	n	Value	n	Value	n	Value	*b* (95% CI)
Depression severity (IDS-SR[b]), mean (SD)							
Baseline	53	45.2 (12.1)	49	41.5 (11.6)	102	43.4 (12.0)	Ref
10 weeks	41	31.9 (12.7)	35	32.0 (17.5)	76	32.0 (15.0)	−3.36 (−8.20, 1.16)
20 weeks	35	30.7 (16.1)	30	27.1 (15.7)	65	29.0 (15.9)	1.50 (−6.22, 9.02)
30 weeks	36	29.5 (17.2)	29	21.1 (15.4)	65	25.8 (16.9)	8.92 (−1.05, 18.50)
QALYs[c] (EQ-5D-3L), mean (SD)							
Baseline	53	0	49	0	102	0	Ref
10 weeks	41	0.09 (0.05)	35	0.10 (0.05)	76	0.10 (0.05)	−0.01 (−0.04, 0.01)
20 weeks	35	0.20 (0.10)	30	0.24 (0.09)	65	0.22 (0.10)	−0.02 (−0.06, 0.03)
30 weeks	36	0.31 (0.16)	29	0.39 (0.13)	65	0.35 (0.15)	−0.03 (−0.10, 0.04)
Reliable change (IDS-SR), n (%)							
Baseline	53	0	49	0	102	0	Ref
10 weeks	41	27 (66)	35	19 (54)	76	46 (61)	0.89 (−1.17, 2.72)
20 weeks	35	21 (60)	30	19 (63)	65	40 (62)	0.07 (−1.69, 2.01)
30 weeks	36	22 (61)	29	23 (79)	65	45 (69)	−0.48 (−2.49, 1.47)
Current depressive episode (MINI-Plus)[d], n (%)							
Baseline	53	52 (98)	49	48 (98)	102	100 (98)	Ref
10 weeks	44	27 (61)	41	18 (44)	85	45 (44)	1.44 (−2.19, 5.07)
20 weeks	44	24 (55)	40	15 (38)	84	39 (38)	1.30 (−2.35, 5.00)
30 weeks	39	15 (39)	37	13 (35)	76	28 (28)	0.32 (−3.35, 4.00)

[a]CBT: cognitive behavioral therapy.

[b]IDS-SR: Inventory of Depressive Symptomatology, self-report version.

[c]MINI-Plus: Mini-International Neuropsychiatric Interview Plus.

[d]QALYs: quality-adjusted life-years.

Table 3. Reliable change in depression severity.

Change at each follow-up	Blended CBT[a]	Standard CBT	Full sample
Week 10	**n=41**	**n=35**	**n=76**
Deterioration, n (%)	1 (2)	3 (9)	4 (5)
No change, n (%)	16 (39)	16 (48)	32 (42)
Improvement, n (%)	21 (68)	10 (29)	31 (41)
Remission, n (%)	3 (7)	6 (17)	9 (12)
Week 20	**n=35**	**n=30**	**n=65**
Deterioration, n (%)	0 (0)	0 (0)	0 (0)
No change, n (%)	16 (46)	13 (43)	29 (45)
Improvement, n (%)	13 (37)	11 (37)	24 (37)
Remission, n (%)	6 (17)	6 (20)	12 (19)
Week 30	**n=36**	**n=29**	**n=65**
Deterioration, n (%)	1 (3)	0 (0)	1 (2)
No change, n (%)	15 (42)	10 (35)	25 (39)
Improvement, n (%)	14 (40)	8 (23)	22 (34)
Remission, n (%)	6 (17)	11 (38)	17 (26)

[a] CBT: cognitive behavioral therapy.

Table 4. Current DSM-IV-TR[a] diagnoses at 30-week follow-up.

Diagnosis	Blended CBT[b] (n=39)	Standard CBT (n=37)	Full sample (N=76)
Any comorbidity, n (%)	17 (44)	18 (49)	35 (46)
Depressive episode, n (%)	15 (39)	13 (27)	28 (37)
Anxiety, n (%)	15 (28)	15 (31)	30 (29)
OCD[c], n (%)	6 (11)	3 (6)	9 (9)
PTSD[d], n (%)	3 (6)	2 (4)	5 (5)
Somatoform, n (%)	4 (8)	4 (8)	8 (8)
Alcohol/drug dependency, n (%)	1 (2)	3 (6)	4 (4)
Comorbid disorders, mean (SD)	1.5 (2.0)	1.3 (1.7)	1.4 (1.9)

[a]DSM-IV: *Diagnostic and Statistical Manual of Mental Disorders* (Fourth Edition).

[b]CBT: Cognitive behavioral therapy.

[c]OCD: obsessive compulsive disorder.

[d]PTSD: posttraumatic stress disorder.

Costs

Before examining cost-effectiveness, costs were explored in separate linear mixed models, controlling for baseline costs. Patients in the blended CBT group reported higher cumulative societal costs on average than patients in standard CBT (overall: b=1410, 95% CI −28.5 to 2776.7; $t_{78.86}$=2.06, P=.04). Part of this effect appeared to be driven by rare and costly medical events. When costs of hospitalization and in-patient psychiatric care were not included in the societal costs in a sensitivity analysis, the between-group difference was reduced (overall: b=1206, 95% CI −300.9 to 2713.3; $t_{76.99}$=1.57, P=.12). Based on the estimated marginal means from the mixed-effects model, the overall societal costs for the full sample at 30 weeks were estimated to be mean €10,075 (95% CI €8086-€12,066).

Regarding average cumulative direct medical costs, both groups reported similar costs (overall: b=55, 95% CI −477.0 to 576.5; $t_{90.19}$=0.22, P=.83). Average direct medical costs at 30 weeks were estimated to be €4535 (95% CI €3789-€5282). Sensitivity analyses showed that for cumulative indirect nonmedical costs (costs associated with absenteeism and presenteeism), patients in the blended CBT group on average reported higher costs compared with standard CBT (overall: b=140, 95% CI −178.4 to 448.4; $t_{76.22}$=0.89, P=.37). Average indirect nonmedical costs at 30 weeks were estimated to be €1413 (95% CI €954-€1872).

Unadjusted (observed) means and outcomes per follow-up assessment are presented in Table 5.

Cost-Effectiveness and Cost Utility

The results from the cost-effectiveness and cost-utility analyses are presented in Table 6. Estimated between-group mean differences in costs, utility, and effects over the full study period were not significant.

Table 5. Unadjusted mean cumulative costs in Euros.

Costs	Blended CBT		Standard CBT		Full sample		Blended CBT vs standard CBT
	n	Mean (SD)	n	Mean (SD)	n	Mean (SD)	*b* (95% CI)
Societal costs (€)							
Baseline	53	2205 (2615)	49	1401 (1761)	102	1819 (891)	Ref
10 weeks	40	5140 (5229)	35	2996 (2608)	75	4140 (2419)	1482 (−114, 2989)
20 weeks	33	9523 (9466)	28	5765 (4974)	61	7798 (5202)	3137 (288, 5980)
30 weeks	30	12,401 (13,198)	24	8493 (7239)	54	10,664 (6316)	3923 (−148, 7963)
Direct medical costs (€)							
Baseline	53	628 (888)	49	411 (389)	102	524 (381)	Ref
10 weeks	40	2363 (2743)	35	1697 (934)	75	2052 (1661)	322 (−257, 920)
20 weeks	33	3834 (3833)	28	3025 (1846)	61	3463 (2610)	308 (−734, 1343)
30 weeks	30	4728 (4504)	24	4.287 (3743)	54	4527 (3398)	25 (−1525, 1512)

[a]CBT: Cognitive behavioral therapy.

Table 6. Results of the cost-effectiveness and cost-utility analyses.

Outcome	Blended CBT[a] vs standard CBT		ICER[b]	Cost-effectiveness plane distribution[c] (%)			
	Difference cost, mean difference (95% CI)	Difference effect, mean or risk difference[d] (95% CI)		NW[e]	SW[f]	NE[g]	SE[h]
Societal perspective							
RCI[i]	1183 (−399, 2765)	0.03 (−.010, .15)	39,433	10.0	—	89.2	0.7
MINI[j]	1183 (−399, 2765)	0.06 (−0.05, 0.19)	19,716	89.2	0.7	10.0	—
QALY[k]	1183 (−399, 2765)	0.01 (−0.03, 0.04)	185,880	34.4	0.6	64.9	0.2
Health care provider perspective							
RCI	−176 (−659, 343)	0.03 (−0.10, 0.15)	−5867	5.2	4.9	20.2	69.7
MINI	−176 (−659, 343)	0.06 (−0.05, 0.19)	−2933	20.2	69.7	5.2	4.9
QALY	−176 (−659, 343)	0.01 (−0.03, 0.04)	−29,333	3.0	31.9	22.3	42.7

[a]CBT: Cognitive behavioral therapy.

[b]ICER: incremental cost-effectiveness ratio.

[c]For plane distribution, NW=more expensive, less effective; SW=less expensive, less effective; NE=more expensive, more effective; SE=less expensive, more effective.

[d]Risk difference: RCI and MINI; Mean: QALY.

[e]NW: northwest.

[f]SW: southwest.

[g]NE: northeast.

[h]SE: southeast.

[i]RCI: reliable change index (based on IDS-SR).

[j]MINI: Mini-International Neuropsychiatric Interview Plus (diagnostic interview diagnosis of a depressive episode).

[k]QALY: quality-adjusted life-year (based on EQ-5D-3L).

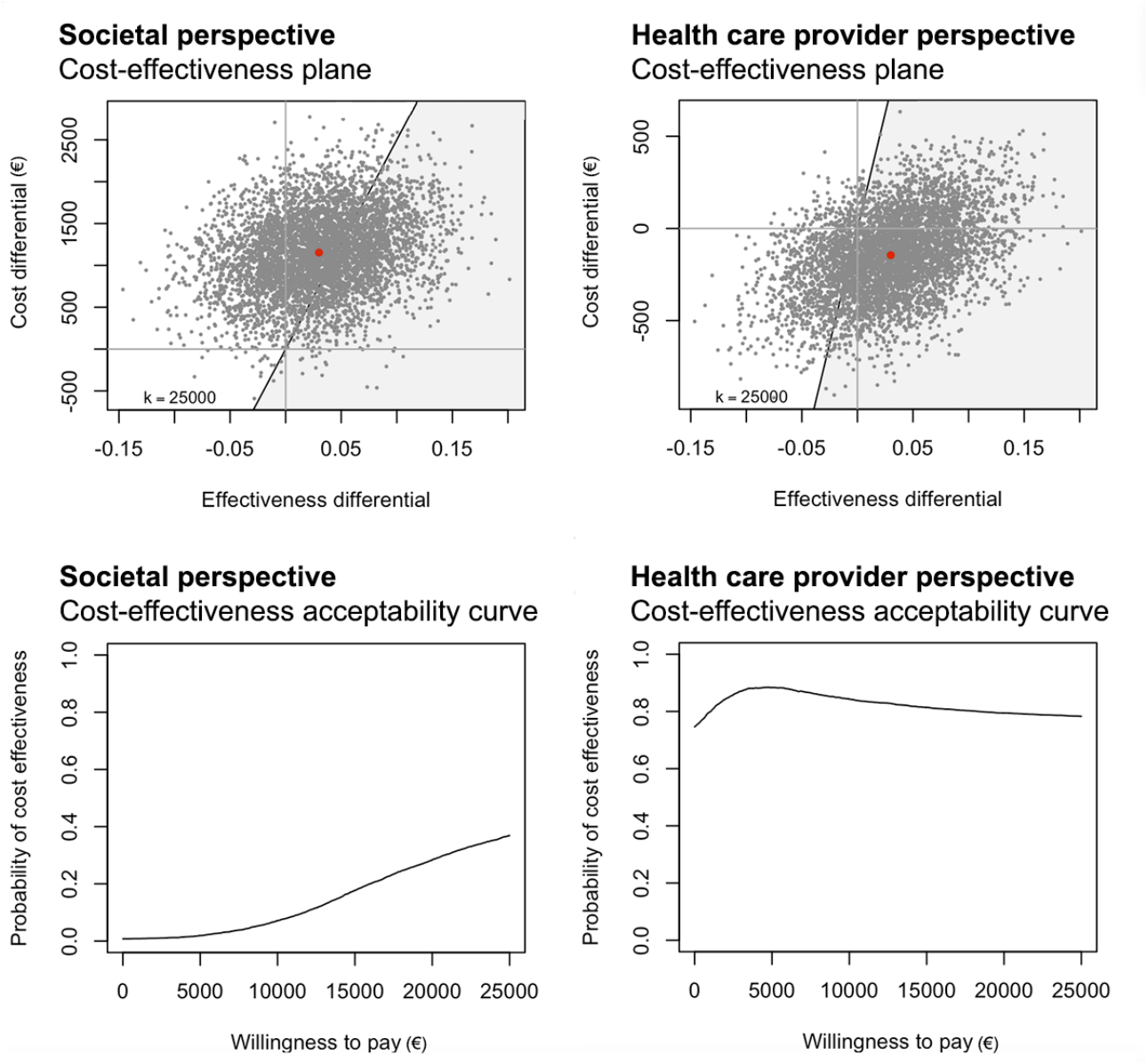
Societal Perspective

For response to treatment, the ICER was 39,433, meaning that an additional treatment response in blended CBT was associated with €39,433 higher costs compared with standard CBT. Of the estimated cost-effect pairs, 89.2% was located in the northeast quadrant (more effective and more expensive), 10% in the northwest quadrant (less effective and more expensive), and 0.7% in the southeast quadrant (more effective and less expensive). The probability of blended CBT being cost-effective compared with standard CBT was 0.01 at a ceiling ratio of €0 per additional response to treatment, and 0.02, 0.18, and 0.28 at ceiling ratios of €5000, €15,000, and €20,000 per response to treatment, respectively. The probability of blended CBT being cost-effective compared with standard CBT rose to 0.37 at a ceiling ratio of €25,000 per additional response to treatment. The cost-effectiveness plane and cost-effectiveness acceptability curve are presented in Figure 2.

For occurrence of depressive episodes, the ICER was 19,716. This means that avoiding an additional depressive episode in blended CBT was associated with €19,716 higher costs compared with standard CBT. The probability of blended CBT being cost-effective compared with standard CBT was 0.01 at a ceiling ratio of €0 per depressive episode. This was the highest possible probability.

For QALYs, the ICER was 185,880, meaning that one extra QALY in blended CBT was associated with €185,880 higher costs compared with standard CBT. The probability of blended CBT being cost-effective compared with standard CBT was 0.01 at a ceiling ratio of €0 per QALY. The probability of blended CBT being cost-effective compared with standard CBT rose to 0.02 at a ceiling ratio of €25,000 per extra QALY.

Figure 2. Cost-effectiveness planes for response to treatment (blended CBT versus standard CBT) from the societal perspective (top left) and health care provider perspective (top right), and cost-effectiveness acceptability curves for response to treatment (blended CBT versus standard CBT) from the societal perspective (bottom left) and health care provider perspective (bottom right). CBT: cognitive behavioral therapy



Health Care Provider Perspective

For response to treatment, the ICER was −5867, meaning that an additional treatment response in blended CBT was associated with €5867 lower costs compared with standard CBT. Of the estimated cost-effect pairs, 69.7% was located in the southeast quadrant (more effective and less expensive), 20.2% in the northeast quadrant (more effective and more expensive), 5.2% in the northwest quadrant (less effective and more expensive), and 4.9% in the southwest quadrant (less effective, less expensive). The probability of blended CBT being cost-effective compared with standard CBT was 0.75 at a ceiling ratio of €0 per additional response to treatment, and 0.80 at a ceiling ratio of €1000 per additional response to treatment. The probability of blended CBT being cost-effective compared with standard CBT rose to 0.88 at a ceiling ratio of €5000 per additional response to treatment. The cost-effectiveness plane and cost-effectiveness acceptability curve are presented in Figure 2.

For occurrence of depressive episodes, the ICER was −2933. This means that an additional depressive episode in blended CBT was associated with €2933 lower costs compared with standard CBT. The probability of blended CBT being cost-effective compared with standard CBT was 0.75 at a ceiling ratio of €0 per depressive episode.

For QALYs, the ICER was −29,333, meaning that one extra QALY in blended CBT was associated with €29,333 lower costs compared with standard CBT. The probability of blended CBT being cost-effective compared with standard CBT was 0.75 at a ceiling ratio of €0 per QALY gained, and 0.81 at €5000 per QALY gained. The probability of blended CBT being cost-effective compared with standard CBT rose to 0.85 at €10,000 per QALY gained.

Discussion

Overview

This pilot RCT focused on depression treatment in Dutch routine specialized mental health care and compared the costs and effects of blended CBT for MDD to standard (evidence-based) face-to-face CBT over a 30-week time frame. The study examined whether blended CBT has the potential to lead to clinical effects that are comparable to the effects of standard CBT, at lower costs. To the best of our knowledge, our study was the first to explore the cost-effectiveness of integrated (rather than add-on) blended depression treatment in routine mental health care and compare this blended treatment to existing, evidence-based, face-to-face CBT.

Principal Findings

The results suggest that blended CBT for depression has the potential to lead to costs and clinical effects that are comparable to the costs and effects associated with standard CBT. In both treatment groups, the severity of depressive symptoms decreased, the probability of having a depression diagnosis lessened, and quality of life improved. When costs and effects were combined, results were mixed. From a societal perspective, which includes productivity losses, blended CBT was not cost-effective compared with standard CBT. From a health care provider perspective, blended CBT had an acceptable probability of being cost-effective compared with standard CBT for treatment response and QALYs (cost utility), but not for depression diagnosis (depressive episodes).

Results show that blended CBT patients received significantly fewer face-to-face sessions than standard CBT patients (10 versus 13 sessions). In line with our expectations, blended CBT also led to shorter treatment duration (19 versus 33 weeks). After 20 weeks of treatment, 51% (27 of 53) of blended CBT patients completed 75% or more of the treatment protocol (at least 14 of 18 sessions in total) versus 5% in the CBT group (3 of 49). After 30 weeks, the percentage of treatment completers rose to 85% for blended CBT and 45% for standard CBT. Combined with therapist time required for online feedback, blended CBT and standard CBT unexpectedly required similar per-patient time investments from therapists (14 versus 13 hours). This was mainly driven by the mean number of sessions in standard CBT, which was considerably lower than expected based on the protocol (13 versus 18 sessions).

Comparison with Other Work

Because this is one of the first studies to compare both costs and effects of blended depression treatment to treatment as usual, the possibility of comparing our findings to those of previous research is limited. The clinical findings appear to be in line with other studies that examined the clinical effectiveness of combined online and face-to-face treatment for depression. However, in this study we defined blended treatment as an integrated approach, replacing face-to-face sessions with online modules rather than a combined approach that adds an online intervention to a face-to-face CBT protocol. For example, Berger and colleagues [33] evaluated the clinical effectiveness of an evidence-based unguided CBT self-help intervention (Deprexis) combined with regular psychotherapy in outpatient specialized depression care compared with regular psychotherapy. The results of their RCT (N=98) showed superiority of the combined treatment [33] over regular treatment for reduced depressive symptoms (via Beck Depression Inventory-II, BDI-II) (Cohen *d*=.51) at posttreatment (12 weeks). In another RCT (N=229), this unguided self-help intervention was added to a psychodynamic psychotherapeutic treatment (treatment as usual) for depressed inpatients and compared with online information plus treatment as usual [35]. After treatment, treatment as usual plus unguided self-help was superior to treatment as usual plus active control (Cohen *d*=.44).

A naturalistic study in routine specialized mental health care by Kenter and colleagues [36] showed comparable clinical effects for blended and standard face-to-face treatment for depressed patients, as measured with Global Assessment of Functioning (GAF) scores (N=3175) [36]. In this study, the combined rather than integrated approach led to longer treatment duration and more treatment sessions (combined total of online and face-to-face sessions), making blended treatment by design more costly than face-to-face treatment [36].

Offering blended treatment in a group format rather than to patients individually could potentially further lower costs associated with treatment. Schuster and colleagues [64] examined change in severity of depressive symptoms (via Center

for Epidemiologic Studies Depression Scale, CES-D) in a combined blended group therapy compared with a waitlist control group. Participants (N=46) were adults with depressive symptoms who were recruited from the general population in Austria. Blended therapy consisted of eight 90-minute group sessions, combined with access to an e-learning platform. Treatment was provided at the University of Salzburg. The group intervention was superior to waitlist (Cohen *d*=.87) after treatment.

Finally, in a recent study by Thase and colleagues [25], an integrated approach was chosen for treating medication-free adults with MDD (N=154). Treatment was provided in two university clinics. In the blended treatment, nine online modules were combined with 12 face-to-face sessions of 25 minutes. Compared with 20 CBT sessions of 50 minutes, blended CBT led to noninferior results (Cohen *d*=.05) on the HAM-D after 16 weeks of treatment.

Limitations

There are a number of factors that need to be considered when interpreting the findings. First, the sample size was smaller than anticipated, which led to less stable probability estimates. This was mainly due to time constraints of conducting the study. The achieved sample is considered adequate for the central aim of the study, which was to guide further development and research in blended CBT. Future studies should focus on hypothesis testing in larger samples, using a noninferiority design.

Second, within this study, patients could be followed for 30 weeks. Although this was considered an acceptable time period to assess changes in outcomes during and shortly after treatment, future studies could include a long-term follow-up (eg, 1 year after treatment). This would provide more insight into the course of depression, treatment trajectories, quality of life, and costs over time. When examining costs, future studies could also include specific costs associated with delivering the interventions, such as the hosting and maintenance of the online treatment platform. Unfortunately, these data were not available in this study.

Third, this sample was characterized by a high level of heterogeneity, both in clinical and cost characteristics. On the one hand, this is a positive aspect because it is representative of the population in routine specialized mental health care. On the other hand, it led to some baseline imbalances between groups. This is most apparent in the average direct medical and nonmedical costs patients reported in the month before baseline, which were higher in the blended CBT group than the standard CBT group. This was primarily caused by the fact that three patients in the blended group reported having been admitted to an in-patient psychiatric ward before baseline assessment versus no patients in the standard CBT group.

The imbalances at baseline increased the uncertainty around the cost and effectiveness estimates, especially at the 30-week follow-up. This is reflected in the analyses concerning treatment response, quality of life, and direct medical costs (health care provider perspective). The models that included time were in favor of the standard CBT group, whereas the models that did not include time were in favor of blended CBT. It is important to keep in mind that all findings were nonsignificant and varied around zero, signifying no differences between groups.

Finally, it should be noted that in the study protocol [41] more outcomes were included than could be reported on in this paper, such as working alliance and depressive cognitions. Information on these outcomes will be provided in forthcoming publications. Next, in the study protocol, it was stated that treatment response would be assessed at 30 weeks. This might lead readers to wonder whether there was a deviation from the planned analyses. However, within the cost-effectiveness framework, the focus lies on cumulative outcomes. Therefore, cumulative group differences were estimated with linear mixed models over the full study period, rather than estimating a group difference at a single time point (30 weeks).

Conclusion

This study is one of the first to examine both costs and effects of integrated blended CBT for depression in specialized mental health care. The results are promising and suggest it is feasible to digitalize part of the therapist-patient interaction, even in the complex patient population that characterizes specialized mental health care. Blended CBT appears to lead to comparable clinical effects as standard CBT, may increase treatment adherence, and could potentially speed up patient flow through the treatment process. From a societal perspective, blended CBT is not considered cost-effective compared with standard CBT. However, there is an acceptable probability that blended CBT is cost-effective from the perspective of the health care provider. Further research in a larger sample seems warranted, which should focus on optimizing the clinical effects of blended treatment as well as the cost-impact.

Acknowledgments

This study was funded by ZonMw (project number 837001007). The authors gratefully acknowledge the contribution of all participants, research assistants, therapists, and all others who contributed to the study. We would like to thank Adriaan Hoogendoorn and Judith Bosmans for their invaluable advice during the data-analysis process.

Authors' Contributions

HR and PvO obtained funding for the study. All authors contributed to the design of the study. HR, PvO, JEW, LCK, and JR developed the intervention. LCK coordinated the recruitment of patients and the data collection. HR, FS, JL, KN, JR, and LK developed the analytic plan. FS, JL, KN, and LCK analyzed the data. LCK wrote the manuscript. All authors read and revised

the manuscript and gave final approval of the version to be published. JR passed away in July 2019, shortly before the manuscript was submitted for the final round of the peer-review process.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Cost categories and prices (€, 2014).
[PDF File (Adobe PDF File), 35 KB - jmir_v21i10e14261_app1.pdf]

Multimedia Appendix 2
Compliance with questionnaires at follow-up.
[PDF File (Adobe PDF File), 21 KB - jmir_v21i10e14261_app2.pdf]

Multimedia Appendix 3
CONSORT-EHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 2515 KB - jmir_v21i10e14261_app3.pdf]

References

1. Cuijpers P, Huibers M, Ebert DD, Koole SL, Andersson G. How much psychotherapy is needed to treat depression? A metaregression analysis. J Affect Disord 2013;149(1-3):1-13. [doi: 10.1016/j.jad.2013.02.030] [Medline: 23528438]
2. Karyotaki E, Smit Y, Holdt Henningsen K, Huibers MJ, Robays J, de Beurs D, et al. Combining pharmacotherapy and psychotherapy or monotherapy for major depression? A meta-analysis on the long-term effects. J Affect Disord 2016 Apr;194:144-152. [doi: 10.1016/j.jad.2016.01.036] [Medline: 26826534]
3. Butler AC, Chapman JE, Forman EM, Beck AT. The empirical status of cognitive-behavioral therapy: a review of meta-analyses. Clin Psychol Rev 2006 Jan;26(1):17-31. [doi: 10.1016/j.cpr.2005.07.003] [Medline: 16199119]
4. Kohn R, Saxena S, Levav I, Saraceno B. The treatment gap in mental health care. Bull World Health Organ 2004 Nov;82(11):858-866 [FREE Full text] [Medline: 15640922]
5. Bremmer F, van Es M. Een analyse van de verwachte kosten en baten van eHealth-blended behandelen en begeleiden. Amersfoort, the Netherlands: GGZ Nederland; 2013.
6. Araya R, Zitko P, Markkula N, Rai D, Jones K. Determinants of access to health care for depression in 49 countries: a multilevel analysis. J Affect Disord 2018 Jul;234:80-88. [doi: 10.1016/j.jad.2018.02.092] [Medline: 29524750]
7. Health Quality Ontario. Psychotherapy for major depressive disorder and generalized anxiety disorder: a health technology assessment. Ont Health Technol Assess Ser 2017;17(15):1-167. [Medline: 29213344]
8. Andersson G. The promise and pitfalls of the internet for cognitive behavioral therapy. BMC Med 2010;8:82 [FREE Full text] [doi: 10.1186/1741-7015-8-82] [Medline: 21138574]
9. Ruwaard J, Kok RN. Wild West eHealth: Time to hold our horses? Eur Health Psychol 2015;17(1):45-49.
10. Emmelkamp PM, David D, Beckers T, Muris P, Cuijpers P, Lutz W, et al. Advancing psychotherapy and evidence-based psychological interventions. Int J Methods Psychiatr Res 2014;23 Suppl 1:58-91. [doi: 10.1002/mpr.1411] [Medline: 24375536]
11. Andrews G, Basu A, Cuijpers P, Craske M, McEvoy P, English C, et al. Computer therapy for the anxiety and depression disorders is effective, acceptable and practical health care: An updated meta-analysis. J Anxiety Disord 2018 Apr;55:70-78 [FREE Full text] [doi: 10.1016/j.janxdis.2018.01.001] [Medline: 29422409]
12. Wright JH, Wright AS, Albano AM, Basco MR, Goldsmith LJ, Raffield T, et al. Computer-assisted cognitive therapy for depression: maintaining efficacy while reducing therapist time. Am J Psychiatry 2005 Jun;162(6):1158-1164. [doi: 10.1176/appi.ajp.162.6.1158] [Medline: 15930065]
13. Richards D, Richardson T. Computer-based psychological treatments for depression: a systematic review and meta-analysis. Clin Psychol Rev 2012 Jun;32(4):329-342. [doi: 10.1016/j.cpr.2012.02.004] [Medline: 22466510]
14. Andersson G, Cuijpers P. Internet-based and other computerized psychological treatments for adult depression: a meta-analysis. Cogn Behav Ther 2009;38(4):196-205. [doi: 10.1080/16506070903318960] [Medline: 20183695]
15. Andersson G, Rozental A, Shafran R, Carlbring P. Long-term effects of internet-supported cognitive behaviour therapy. Expert Rev Neurother 2018 Jan;18(1):21-28. [doi: 10.1080/14737175.2018.1400381] [Medline: 29094622]
16. Webb CA, Rosso IM, Rauch SL. Internet-based cognitive-behavioral therapy for depression: current progress and future directions. Harv Rev Psychiatry 2017;25(3):114-122 [FREE Full text] [doi: 10.1097/HRP.0000000000000139] [Medline: 28475503]
17. Carlbring P, Andersson G, Cuijpers P, Riper H, Hedman-Lagerlöf E. Internet-based vs face-to-face cognitive behavior therapy for psychiatric and somatic disorders: An updated systematic review and meta-analysis. Cogn Behav Ther 2018;47(1):1-18. [doi: 10.1080/16506073.2017.1401115] [Medline: 29215315]

18. Kolovos S, van Dongen JM, Riper H, Buntrock C, Cuijpers P, Ebert DD, et al. Cost effectiveness of guided Internet-based interventions for depression in comparison with control conditions: an individual-participant data meta-analysis. Depress Anxiety 2018;35(3):209-219 [FREE Full text] [doi: 10.1002/da.22714] [Medline: 29329486]
19. Cuijpers P, Riper H, Andersson G. Internet-based treatment of depression. Curr Opin Psychol 2015;4(1):131-135 [FREE Full text] [doi: 10.1016/j.copsyc.2014.12.026]
20. Schueller SM, Tomasino KN, Mohr DC. Integrating human support into behavioral intervention technologies: the efficiency model of support. Clin Psychol Sci Pract 2016;24(1):27-45. [doi: 10.1111/cpsp.12173]
21. Wentzel J, van der Vaart R, Bohlmeijer ET, van Gemert-Pijnen JE. Mixing online and face-to-face therapy: how to benefit from blended care in mental health care. JMIR Ment Health 2016;3(1):e9 [FREE Full text] [doi: 10.2196/mental.4534] [Medline: 26860537]
22. Kooistra LC, Ruwaard J, Wiersma JE, van Oppen P, van der Vaart R, van Gemert-Pijnen JE, et al. Development and initial evaluation of blended cognitive behavioural treatment for major depression in routine specialized mental health care. Internet Interv 2016 May;4:61-71 [FREE Full text] [doi: 10.1016/j.invent.2016.01.003] [Medline: 30135791]
23. Erbe D, Eichert HC, Riper H, Ebert DD. Blending face-to-face and Internet-based interventions for the treatment of mental disorders in adults: systematic review. J Med Internet Res 2017;19(9):e306 [FREE Full text] [doi: 10.2196/jmir.6588] [Medline: 28916506]
24. Andersson G, Cuijpers P. Pros and cons of online cognitive-behavioural therapy. Br J Psychiatry 2008 Oct;193(4):270-271. [doi: 10.1192/bjp.bp.108.054080] [Medline: 18827286]
25. Thase ME, Wright JH, Eells TD, Barrett MS, Wisniewski SR, Balasubramani GK, et al. Improving the efficiency of psychotherapy for depression: computer-assisted versus standard CBT. Am J Psychiatry 2018;175(3):242-250. [doi: 10.1176/appi.ajp.2017.17010089] [Medline: 28969439]
26. Kleiboer A, Smit J, Bosmans J, Ruwaard J, Andersson G, Topooco N, et al. European COMPARative Effectiveness research on blended depression treatment versus treatment-as-usual (E-COMPARED): study protocol for a randomized controlled, non-inferiority trial in eight European countries. Trials 2016;17(1):387 [FREE Full text] [doi: 10.1186/s13063-016-1511-1] [Medline: 27488181]
27. Romijn G, Riper H, Kok R, Donker T, Goorden M, Hakkaart-van Roijen L, et al. Cost-effectiveness of blended vs face-to-face cognitive behavioural therapy for severe anxiety disorders: study protocol of a randomized controlled trial. BMC Psychiatry 2015;15:311 [FREE Full text] [doi: 10.1186/s12888-015-0697-1] [Medline: 26651478]
28. Vis C, Kleiboer A, Prior R, Bønes E, Cavallo M, Clark S, et al. Implementing and up-scaling evidence-based eMental health in Europe: the study protocol for the MasterMind project. Internet Interventions 2015;2(4):399-409. [doi: 10.1016/j.invent.2015.10.002]
29. Mol M, Dozeman E, van Schaik DJ, Vis CP, Riper H, Smit JH. The therapist's role in the implementation of internet-based cognitive behavioural therapy for patients with depression: study protocol. BMC Psychiatry 2016;16(1):338 [FREE Full text] [doi: 10.1186/s12888-016-1045-9] [Medline: 27716108]
30. Hickie IB, Davenport TA, Luscombe GM, Moore M, Griffiths KM, Christensen H. Practitioner-supported delivery of internet-based cognitive behaviour therapy: evaluation of the feasibility of conducting a cluster randomised trial. Med J Aust 2010;192(11 Suppl):S31-S35. [Medline: 20528705]
31. Høifødt RS, Lillevoll KR, Griffiths KM, Wilsgaard T, Eisemann M, Waterloo K, et al. The clinical effectiveness of web-based cognitive behavioral therapy with face-to-face therapist support for depressed primary care patients: randomized controlled trial. J Med Internet Res 2013;15(8):e153 [FREE Full text] [doi: 10.2196/jmir.2714] [Medline: 23916965]
32. Kessler D, Lewis G, Kaur S, Wiles N, King M, Weich S, et al. Therapist-delivered Internet psychotherapy for depression in primary care: a randomised controlled trial. Lancet 2009 Aug 22;374(9690):628-634. [doi: 10.1016/S0140-6736(09)61257-5] [Medline: 19700005]
33. Berger T, Krieger T, Sude K, Meyer B, Maercker A. Evaluating an e-mental health program. J Affect Disord 2018;227:455-462. [doi: 10.1016/j.jad.2017.11.021] [Medline: 29154168]
34. Meyer B, Berger T, Caspar F, Beevers CG, Andersson G, Weiss M. Effectiveness of a novel integrative online treatment for depression (Deprexis): randomized controlled trial. J Med Internet Res 2009;11(2):e15 [FREE Full text] [doi: 10.2196/jmir.1151] [Medline: 19632969]
35. Zwerenz R, Becker J, Knickenberg RJ, Siepmann M, Hagen K, Beutel ME. Online self-help as an add-on to inpatient psychotherapy: efficacy of a new blended treatment approach. Psychother Psychosom 2017;86(6):341-350. [doi: 10.1159/000481177] [Medline: 29131090]
36. Kenter RM, van de Ven PM, Cuijpers P, Koole G, Niamat S, Gerrits RS, et al. Costs and effects of Internet cognitive behavioral treatment blended with face-to-face treatment: results from a naturalistic study. Internet Interventions 2015;2(1):77-83. [doi: 10.1016/j.invent.2015.01.001]
37. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders, 4th Edition, Text Revision (DSM-IV-TR). Washington, DC: American Psychiatric Publishing; 2000.
38. Spijker J, Bockting CL, Meeuwissen JA, van Vliet IM, Emmelkamp PM, Hermens ML, et al. Multidisciplinaire richtlijn depressie (derde revisie). Richtlijn voor de diagnostiek, behandeling en begeleiding van volwassen patiënten met een depressieve stoornis. Utrecht, the Netherlands: Trimbos-instituut; 2013.

39. Sheehan DV, Lecrubier Y, Sheehan KH, Amorim P, Janavs J, Weiller E, et al. The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. J Clin Psychiatry 1998;59 Suppl 20:22-33;quiz 34. [Medline: 9881538]
40. van Vliet IM, de Beurs E. Het Mini Internationaal Neuropsychiatrisch Interview (MINI): een kort gestructureerd diagnostisch psychiatrisch interview voor DSM-IV en ICD-10-stoornissen. Tijdschrift voor de Psychiatrie 2007;49(6):197-393.
41. Kooistra LC, Wiersma JE, Ruwaard J, van Oppen P, Smit F, Lokkerbol J, et al. Blended vs. face-to-face cognitive behavioural treatment for major depression in specialized mental health care: study protocol of a randomized controlled cost-effectiveness trial. BMC Psychiatry 2014;14:290 [FREE Full text] [doi: 10.1186/s12888-014-0290-z] [Medline: 25326035]
42. Bockting CL, Huibers MJ. Protocollaire behandeling van patienten met een depressieve stoornis. In: Keijsers GPJ GJ, van Minnen A, Hoogduin CA, editors. Protocollaire behandelingen voor volwassenen met psychische klachten deel 1. Amsterdam, the Netherlands: Uitgeverij Boom; 2011:251-288.
43. Beck JS. Cognitive Behavior Therapy: Basics and Beyond. New York: Guilford Press; 2011.
44. Beck AT, Rush AJ, Shaw BF, Emery G. Cognitive Therapy of Depression. New York: Guilford Press; 1980.
45. Rush AJ, Trivedi MH, Ibrahim HM, Carmody TJ, Arnow B, Klein DN, et al. The 16-Item Quick Inventory of Depressive Symptomatology (QIDS), clinician rating (QIDS-C), and self-report (QIDS-SR): a psychometric evaluation in patients with chronic major depression. Biol Psychiatry 2003 Sep 01;54(5):573-583. [Medline: 12946886]
46. Rush AJ, Carmody T, Reimitz P. The Inventory of Depressive Symptomatology (IDS): clinician (IDS - C) and self - report (IDS - SR) ratings of depressive symptoms. Int J Method Psychiat Res 2000;9(2):45-59. [doi: 10.1002/mpr.79]
47. Cusin C, Yang H, Yeung A, Fava M. Rating scales for depression. In: Bear L, Blais MA, editors. Handbook of Clinical Rating Scales and Assessment in Psychiatry and Mental Health. Totowa, NJ: Humana Press; 2019.
48. Rush AJ, Gullion CM, Basco MR, Jarrett RB, Trivedi MH. The Inventory of Depressive Symptomatology (IDS): psychometric properties. Psychol Med 1996 May;26(3):477-486. [Medline: 8733206]
49. Trivedi MH, Rush AJ, Ibrahim HM, Carmody TJ, Biggs MM, Suppes T, et al. The Inventory of Depressive Symptomatology, Clinician Rating (IDS-C) and Self-Report (IDS-SR), and the Quick Inventory of Depressive Symptomatology, Clinician Rating (QIDS-C) and Self-Report (QIDS-SR) in public sector patients with mood disorders: a psychometric evaluation. Psychol Med 2004;34(1):73-82. [doi: 10.1017/s0033291703001107] [Medline: 14971628]
50. Jacobson NS, Truax P. Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. In: Kazdin AE, editor. Methodological Issues and Strategies in Clinical Research. Washington, DC: American Psychological Association; Feb 1991:631-648.
51. Trivedi MH, Rush AJ, Wisniewski SR, Nierenberg AA, Warden D, Ritz L, STAR*D study team. Evaluation of outcomes with citalopram for depression using measurement-based care in STAR*D: implications for clinical practice. Am J Psychiatry 2006 Jan;163(1):28-40. [doi: 10.1176/appi.ajp.163.1.28] [Medline: 16390886]
52. EuroQol Group. EuroQol--a new facility for the measurement of health-related quality of life. Health Policy 1990 Dec;16(3):199-208. [Medline: 10109801]
53. Lamers LM, McDonnell J, Stalmeier PF, Krabbe PF, Busschbach JJ. The Dutch tariff: results and arguments for an effective design for national EQ-5D valuation studies. Health Econ 2006;15(10):1121-1132. [doi: 10.1002/hec.1124] [Medline: 16786549]
54. Hakkaart-van Roijen L, van Straten A, Tiemens B, Donker M. Trimbos/iMTA questionnaire for costs associated with psychiatric illness (TIC-P). Rotterdam, the Netherlands: Institute of Medical Technology Assessment (iMTA); 2002.
55. Hakkaart-van Roijen L, van der Linden N, Bouwmans C, Kanters T, Tan S. Kostenhandleiding: methodologie van kostenonderzoek en referentieprijzen voor economische evaluaties in de gezondheidszorg. Rotterdam, the Netherlands: Institute of Medical Technology Assessment (iMTA); 2016.
56. Hakkaart-Van Roijen L, Tan SS, Bouwmans CA. Handleiding voor kostenonderzoek--methoden en standaard kostprijzen voor economische evaluaties in de gezondheidszorg. Rotterdam, the Netherlands: Instituut voor Medical Technology Assessment (iMTA), Erasmus Universiteit Rotterdam; 2010.
57. Z-index. The Hague, the Netherlands: Zorginstituut Nederland; 2016.
58. Hesser H. Modeling individual differences in randomized experiments using growth models: recommendations for design, statistical analysis and reporting of results of internet interventions. Internet Interventions 2015 May;2(2):110-120. [doi: 10.1016/j.invent.2015.02.003]
59. White IR, Carpenter J, Horton NJ. Including all individuals is not enough: lessons for intention-to-treat analysis. Clinical Trials 2012 Jul 02;9(4):396-407. [doi: 10.1177/1740774512450098]
60. Bates D, Mächler M, Bolker B, Walker S. Fitting linear mixed-effects models using lme4. J Stat Soft 2015;67(1):1-48. [doi: 10.18637/jss.v067.i01]
61. R Core Team. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2018.
62. Fenwick E, O'Brien BJ, Briggs A. Cost-effectiveness acceptability curves--facts, fallacies and frequently asked questions. Health Econ 2004;13(5):405-415. [doi: 10.1002/hec.903] [Medline: 15127421]
63. Hakkaart-van Roijen L, van Straten A, Maiwenn A, Rutten F, Donker M. Cost-utility of brief psychological treatment for depression and anxiety. Br J Psychiatry 2006;188:323-329. [doi: 10.1192/bjp.188.4.323] [Medline: 16582058]

64. Schuster R, Leitner I, Carlbring P, Laireiter A. Exploring blended group interventions for depression: randomised controlled feasibility study of a blended computer- and multimedia-supported psychoeducational group intervention for adults with depressive symptoms. Internet Interv 2017;8:63-71 [FREE Full text] [doi: 10.1016/j.invent.2017.04.001] [Medline: 30135830]

Abbreviations

CBT: cognitive behavioral therapy
ICER: Incremental Cost-Effectiveness Ratio
IDS-SR: Inventory of Depressive Symptomatology Self-Report
MDD: major depressive disorder
MINI-Plus: Mini-International Neuropsychiatric Interview Plus
QALY: quality-adjusted life-year
RCI: reliable change index
TiC-P: Trimbos and iMTA questionnaire on Costs associated with Psychiatric illness

Edited by G Eysenbach; submitted 04.04.19; peer-reviewed by S Kolovos, R Schuster; comments to author 11.05.19; revised version received 25.07.19; accepted 18.08.19; published 29.10.19.

Please cite as:
Kooistra LC, Wiersma JE, Ruwaard J, Neijenhuijs K, Lokkerbol J, van Oppen P, Smit F, Riper H
Cost and Effectiveness of Blended Versus Standard Cognitive Behavioral Therapy for Outpatients With Depression in Routine Specialized Mental Health Care: Pilot Randomized Controlled Trial
J Med Internet Res 2019;21(10):e14261
URL: http://www.jmir.org/2019/10/e14261/
doi:10.2196/14261
PMID:31663855

Original Paper

Growing Disparities in Patient-Provider Messaging: Trend Analysis Before and After Supportive Policy

Nicole Senft[1], PhD; Evan Butler[2], BSc; Jordan Everson[2], PhD

[1]Department of Medicine, Vanderbilt University Medical Center, Nashville, TN, United States

[2]Department of Health Policy, Vanderbilt University School of Medicine, Nashville, TN, United States

Corresponding Author:
Jordan Everson, PhD
Department of Health Policy
Vanderbilt University School of Medicine
2525 West End Avenue
Suite 1275
Nashville, TN, 37203
United States
Phone: 1 16158756207
Email: jordan.everson@vanderbilt.edu

Abstract

Background: Public policy introduced since 2011 has supported provider adoption of electronic medical records (EMRs) and patient-provider messaging, primarily through financial incentives. It is unclear how disparities in patients' use of incentivized electronic health (eHealth) tools, like patient-provider messaging, have changed over time relative to disparities in use of eHealth tools that were not directly incentivized.

Objective: This study examines trends in eHealth disparities before and after the introduction of US federal financial incentives. We compare rates of patient-provider messaging, which was directly incentivized, with rates of looking for health information on the Web, which was not directly incentivized.

Methods: We used nationally representative Health Information National Trends Survey data from 2003 to 2018 (N=37,300) to describe disparities in patient-provider messaging and looking for health information on the Web. We first reported the percentage of individuals across education and racial and ethnic groups who reported using these tools in each survey year and compared changes in unadjusted disparities during preincentive (2003-2011) and postincentive (2011-2018) periods. Using multivariable linear probability models, we then examined adjusted effects of education and race and ethnicity in 3 periods—preincentive (2003-2005), early incentive (2011-2013), and postincentive (2017-2018)—controlling for sociodemographic and health factors. In the postincentive period, an additional model tested whether internet adoption, provider access, or providers' use of EMRs explained disparities.

Results: From 2003 to 2018, overall rates of provider messaging increased from 4% to 36%. The gap in provider messaging between the highest and lowest education groups increased by 10 percentage points preincentive (P<.001) and 22 additional points postincentive (P<.001). The gap between Hispanics and non-Hispanic whites increased by 3.2 points preincentive (P=.42) and 11 additional points postincentive (P=.01). Trends for blacks resembled those for Hispanics, whereas trends for Asians resembled those for non-Hispanic whites. In contrast, education-based disparities in looking for health information on the Web (which was not directly incentivized) did not significantly change in preincentive or postincentive periods, whereas racial disparities narrowed by 15 percentage points preincentive (P=.008) and did not significantly change postincentive. After adjusting for other sociodemographic and health factors, observed associations were similar to unadjusted associations, though smaller in magnitude. Including internet adoption, provider access, and providers' use of EMRs in the postincentive model attenuated, but did not eliminate, education-based disparities in provider messaging and looking for health information on the Web. Racial and ethnic disparities were no longer statistically significant in adjusted models.

Conclusions: Disparities in provider messaging widened over time, particularly following federal financial incentives. Meanwhile, disparities in looking for health information on the Web remained stable or narrowed. Incentives may have disproportionately benefited socioeconomically advantaged groups. Future policy could address disparities by incentivizing providers treating these populations to adopt messaging capabilities and encouraging patients' use of messaging.

(J Med Internet Res 2019;21(10):e14976) doi:10.2196/14976

KEYWORDS

eHealth; policy; communication; secure messaging; disparities; socioeconomic factors; inequality

Introduction

Background

The use of Web-based tools to access health information and health services (electronic health, eHealth) has increased along with the widespread dissemination of the internet [1,2]. However, since 2000, observers have noted a persistent *digital divide* in the diffusion of eHealth tools across segments of the US population, with members of traditionally underserved groups (lower socioeconomic status, older, and racial and ethnic minorities) being less likely to engage in many eHealth activities [3-7]. Over the past decade, several federally supported initiatives, such as regional extension centers and state cooperatives, were implemented to accelerate the spread of some eHealth tools, including Web-based communication between patients and providers. The largest supportive policy was direct financial incentives for *Meaningful Use* (MU) of eHealth records, which began in 2011 and was redesigned as the Promoting Interoperability program within the Merit-Based Incentive Payment System in 2017. Other eHealth tools, such as those facilitating access to Web-based health information that is independent of health care providers, have spread without direct health policy intervention. It is not clear how the digital divide in the use of eHealth tools financially incentivized by federal policy has changed since the enactment of supportive policy compared with the digital divide in the use of eHealth tools that were not directly incentivized by policy makers.

Trends through 2014 suggested that people with higher incomes and education remained more likely to use the internet to look up health information and to communicate with their providers, with disparities based on race sometimes becoming statistically nonsignificant in adjusted models [2,8,9]. Throughout this period, group differences in internet and home computer access at least partly explained disparities in eHealth use [5,10]. Rapid dissemination of the internet, encouraged in part by public support through programs such as the Broadband Opportunities Technology Program, have reduced racial and socioeconomic disparities in internet access and use in recent years [11]. The digital divide in overall eHealth use may have similarly decreased, although these trends have not been examined in recent data. It is also possible that disparities in eHealth use persist, perhaps because of differences in eHealth literacy or perceived benefits and concerns of using eHealth [4,12,13]. Furthermore, disparities in use of eHealth tools to communicate with providers may be uniquely driven by differences in access to health care providers [14]. Individuals who have not seen any provider are not likely to communicate with one on the Web [15], and racial and socioeconomic disparities in access to providers persist despite increased insurance coverage following the Affordable Care Act [12].

Beyond differences in provider access, disparities may be driven by uneven adoption of functional electronic medical records (EMRs) among providers [16,17]. The MU program was designed as an all-or-nothing incentive in which providers either qualified in a year and received a substantial payment or did not receive any incentive payment. MU criteria became progressively more difficult: In 2011, providers were encouraged to develop the ability to send patients reminder messages to receive incentive payments. By 2015, providers were required to send secure messages to at least 5% of unique patients to receive incentives. Only 62% of physicians were able to attest to MU by 2016, and high-resource physician offices (which are more likely to treat high-resource patients) were more likely to attest to MU than low-resource physician offices [16,18,19]. Therefore, it is possible that incentives encouraging electronic communication between patients and providers have unintentionally widened disparities in its adoption and use by inadequately addressing access to providers and, in particular, to providers with highly usable EMRs who encourage eHealth communication [20].

Objectives

In this study, we characterized trends in disparities in 2 eHealth technologies from 2003 to 2018 to examine whether the digital divide persists and whether financial incentives contributed to narrowing or widening that divide. Specifically, we described the use of 1 incentivized technology (communicating with providers via messaging) and 1 technology that was not directly incentivized (looking for health information on the Web) over time and across socioeconomic strata and racial groups. Finally, we investigated whether internet adoption, health care access, and providers' use of EMRs explain the digital divide in recent years.

Methods

Data

We used data from the National Cancer Institute's Health Information National Trends Survey (HINTS), a cross-sectional, nationally representative survey of noninstitutionalized adults in the United States. HINTS was developed to monitor changes in health communication and information technology. We used data from 8 iterations of HINTS, administered in 2003 (HINTS 1), 2005 (HINTS 2), 2008 (HINTS 3), 2011 (HINTS 4, cycle 1), 2013 (HINTS 4, cycle 3), 2015 (HINTS-FDA), 2017 (HINTS 5, cycle 1), and 2018 (HINTS 5, cycle 2). Detailed information on each iteration's sampling methodology, data collection, and response rates are published by the National Cancer Institute [21].

Population

The full sample included 37,300 individual-year responses from 2003 to 2018. As different questions were included in different years, we excluded observations with missing data on an analysis-by-analysis basis, as described in the Analysis section.

Dependent Variables: Electronic Health Use

Use of provider messaging was assessed by asking participants whether, in the past 12 months, they had "used e-mail or the Internet to communicate with a doctor or a doctor's office." This item was not asked in 2015 (HINTS-FDA). Looking for health information on the Web was assessed by asking participants whether, in the past 12 months, they had used the internet to "look for health or medical information for yourself." This item was not asked in 2008 (HINTS 3). Response options for both items were *yes/no*. Before 2017, only individuals who reported using the internet (n=20,445) were asked to respond to these questions. Noninternet users (n=9994) in these survey years, who were not asked these items, were recoded as responding "no" to them, as it is reasonable to assume noninternet users were not using the internet to message their providers or look up health information.

Independent Variables: Sociodemographic and Health-Related Variables

We defined racial and ethnic groups and socioeconomic strata using 2 measures included on all years of HINTS data. To measure race and ethnicity, we defined 6 different racial and ethnic groups: Hispanic, non-Hispanic white, black, Asian, other (American Indian/Native Alaskan/Native Hawaiian/Pacific Islander), and multiracial. We defined socioeconomic strata by the level of education, categorized as 4 levels in all years: less than high school, high school graduate or General Education Diploma, some college or technical school, and college graduate or greater.

We included several additional demographic and health-related variables in multivariable models. Demographic variables included household income (<US $20,000, US $20,000- US $34,999, US $35,000- US $49,999, US $50,000- US $74,999, US $75,000- US $99,999, and > US $100,000), sex (male and female), age (18-34 years, 35-49 years, 50-64 years, 65-74 years, and >75 years), and marital status (married, living as married/member of an unmarried couple, divorced, widowed, separated, and single). Measurement of household income in 2003 differed from other survey years: the lowest income group was less than US $25,000 and the highest was greater than or equal to US $75,000. Health-related variables included health insurance coverage (yes or no) and general health (excellent, very good, good, fair, and poor).

Explanatory Factors

We explored 3 potential explanations for observed health disparities in 2017 to 2018: internet adoption, access to a health care provider, and having a provider that uses an EMR. Internet adoption was assessed by a single item: "Do you ever go on-line to access the Internet or World Wide Web, or to send and receive e-mail?" (yes/no). Access to a health care provider was assessed by the item, "In the past 12 months, not counting times you went to an emergency room, how many times did you go to a doctor, nurse, or other health professional to get care for yourself?" This item was dichotomized to reflect those who had versus those who had not visited a health care provider in the past year. Finally, provider use of an EMR was defined using the item, "Do any of your doctors or other health care providers maintain your medical records in a computerized system?" (yes/no).

Analysis

We first described the overall rates of provider messaging and looking for health information on the Web over time by plotting the percentage of all individuals who reported using these eHealth tools in each survey year. Analyses included all nonmissing responses to provider messaging (n=32,742, average responses per year=4677; minimum=2905, maximum=7612) and looking for health information on the Web (n=28,663, average responses per year=4090; minimum=2885, maximum=6350).

Next, we described disparities over time in provider messaging and looking for health information on the Web by plotting the percentage of individuals across levels of education and by racial and ethnic group who reported using these tools in each survey year. Multimedia Appendix 1 additionally plots the use of these tools across income levels.

We then compared the unadjusted change in magnitude of disparities in eHealth use by education and race and ethnicity during the preincentive period (2003 vs 2011) and postincentive period (2011 vs 2018). Linear regression models included an indicator for year, education level or race and ethnicity, and the interaction between year and education or race and ethnicity. We added and subtracted regression coefficients using linear combinations to generate mean differences and SEs of differences across groups and years. Two racial and ethnic groups (Native American/Pacific Islander and Multiracial) were excluded from this analysis because they did not have at least 100 observations in each survey year. Respondents were excluded from analyses if they were missing data on education (3%) or race (7%).

We generated adjusted estimates of the associations of education and race and ethnicity with provider messaging and looking for health information on the Web using multivariable linear probability models. Analyses included both education and race and ethnicity and were adjusted for household income, sex, age, marital status, insurance coverage, and general health. To examine changing adjusted associations over time, we replicated models in 3 separate periods: before major public investment (2003-2005), the first years of public support (2011-2013), and recent years (2017-2018). In constructing these periods, we excluded the 2008 and 2015 survey years in which only 1 dependent variable was included in the survey instrument. Finally, in 2017 to 2018, an additional regression model added internet adoption, access to health care providers, and providers' use of an EMR to the multivariable model described above to test whether these factors explained the disparities that persisted in 2017 to 2018. Only cases with complete data were used in analyses; sample sizes for each regression analysis can be found in Table 1. Survey weights were applied in all analyses, which were conducted using Stata 15 MP by StataCorp, College Station, TX.

Table 1. The associations of education and race with electronic health use in preincentive (2003-2005), early incentive (2011-2013), and postincentive (2017-2018) periods. Analyses are adjusted for individuals' income, age, gender, marital status, insurance status, and health status. Linear probability models were generated using complete case analyses. Unweighted sample sizes for each model are provided in parentheses. Survey weights were used to generate means reflective of the US population. *P* values were created using jackknife SEs. Full regression results and SEs are available in Multimedia Appendix 2.

Independent variables	Provider messaging				Looking for health information on the Web			
	2003-2005 (n=9954)	2011-2013 (n=5292)	2017-2018A (n=5305)	2017-2018B[a] (n=5294)	2003-2005 (n=9950)	2011-2013 (n=5333)	2017-2018A (n=5317)	2017-2018B (n=5306)
Education level (Reference: Less than high school)								
High school graduate	−0.001	−0.01	0.05	0.03	0.08[b]	0.09	0.07	0.02
Some college	0.03[b]	0.09[b]	0.14[b]	0.11[c]	0.22[b]	0.27[b]	0.21[b]	0.13
College graduate	0.06[b]	0.13[b]	0.24[b]	0.18[c]	0.34[b]	0.35[b]	0.28[b]	0.19[b]
Race (Reference: non-Hispanic white)								
Hispanic	−0.02	−0.002	−0.05[a]	−0.03	−0.11[b]	−0.05	−0.05	−0.01
Black	−0.00	−0.01	0.02	−0.03	−0.07[b]	−0.02	−0.03	−0.01
Asian	−0.02	0.05	0.03	−0.06	−0.02	−0.05	−0.03	0.01
Multiple races selected	0.01	0.03	0.02	−0.03	0.04	−0.14	0.09[c]	0.08
Other	0.01	−0.003	−0.01	−0.06	−0.10[c]	−0.22[c]	−0.26	−0.26
Explanatory factors								
Uses internet	—[d]	—	—	0.14[b]	—	—	—	0.39[b]
Seen physician in previous 12 months	—	—	—	0.11[b]	—	—	—	0.10[b]
Physician uses EMR[e] (Reference: No)								
Yes	—	—	—	0.20[b]	—	—	—	0.05
Do not know	—	—	—	0.03	—	—	—	0.02
Constant	0.01	0.05	0.06	−0.21	0.21	0.48	0.52	0.19

[a]An additional model (2017-2018B) included potential explanatory factors: internet adoption, provider access, and providers' use of EMRs.

[b]*P*<.01.

[c]*P*<.05.

[d]Not applicable.

[d]EMR: electronic medical record.

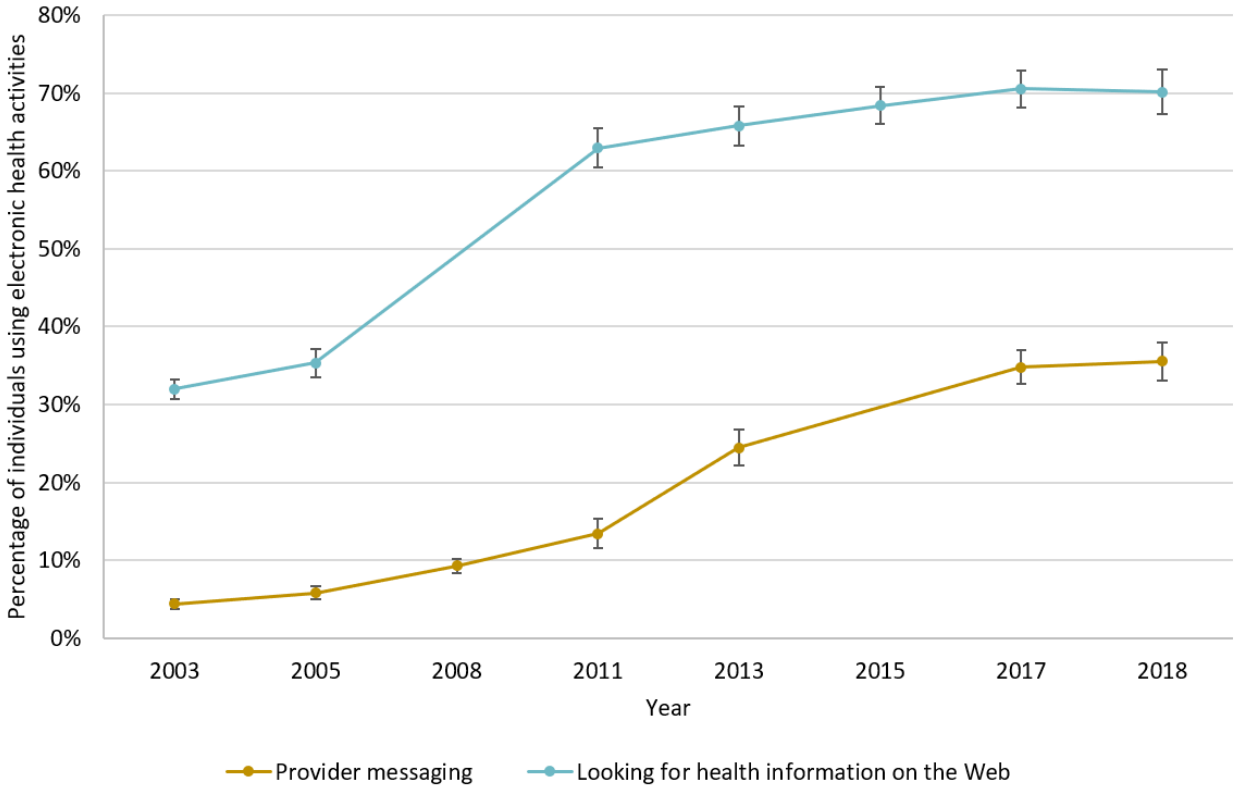
Results

Overall Trends

Provider Messaging

The population-weighted percentage of individuals using provider messaging increased by 32 percentage points, from 4.4% in 2003 to 36% in 2018 (Figure 1). Growth was relatively slow during the first years of the study period, increasing by 9.0 percentage points (from 4.4% to 13%) between 2003 and 2011. Growth was more rapid in later years, after the enactment of relevant public policies starting in 2011, increasing by 22 percentage points between 2011 and 2018 (from 13% to 36%).

Figure 1. Prevalence of electronic health use, 2003 to 2018. The sample for provider messaging includes 32,742 total responses (average 4677 per year), and the sample for looking for health information on the Web includes 28,663 total responses (average 4090 per year). Survey weights were used to generate means reflective of the US population. Bars represent 95% CIs generated using jackknife SEs.



Looking for Health Information on the Web

The percentage of individuals that reported looking for health information on the Web increased by 38 percentage points, from 32% in 2003 to 70% in 2018. In contrast to provider messaging, growth was rapid during the first years of the study period and slowed in later years. Between 2003 and 2011, the percentage increased by 31 percentage points (from 32% to 63%) and then increased by only 7.2 percentage points between 2011 and 2018 (from 63% to 70%).

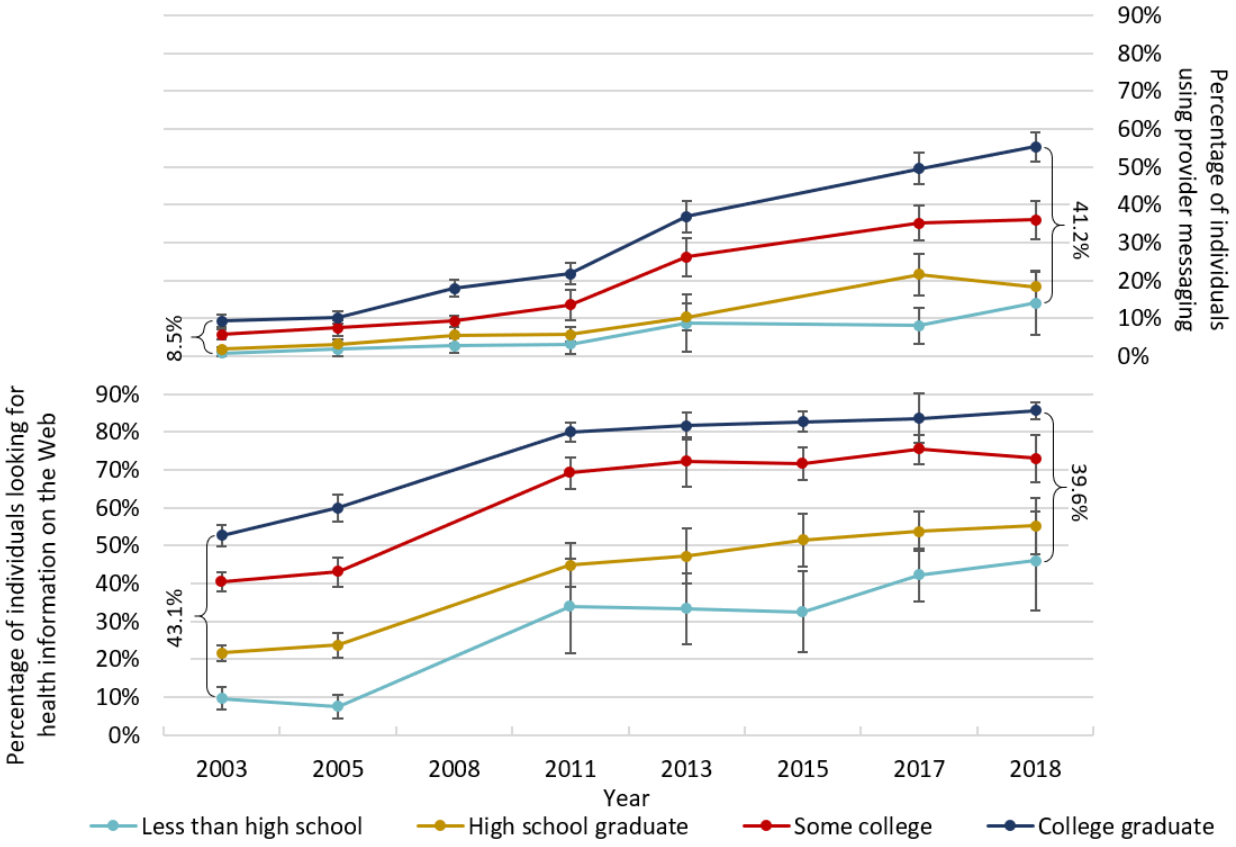
Trends in Disparities

Education

Reported use of provider messaging increased across all education groups between 2003 and 2018 (Figure 2, top panel), although growth was slower among individuals with lower levels of education. Among individuals who did not complete high school, rates of provider messaging increased by only 14 percentage points, with most of this increase (11 points) occurring after 2011. Over the same period, there was a 46 percentage point increase among college graduates, again with most of the change (34 points) occurring after 2011. The gap between the highest and lowest education groups increased by 10 percentage points between 2003 and 2011 ($P<.001$) and an additional 22 percentage points between 2011 and 2018 ($P<.001$).

Reported use of the internet to look for health information also increased across all education groups (Figure 2, bottom panel). In contrast to provider messaging, the increase in looking for health information on the Web was similar across education levels. From 2003 to 2018, reported rates increased by 36 percentage points among individuals who did not complete high school, compared with 33 percentage points among those with a college degree. For both groups, the majority of this increase occurred in the preincentive period; from 2003 to 2011, rates increased by 27 percentage points among individuals who did not complete high school and 24 percentage points among college graduates. Therefore, the gap between the highest and lowest education groups increased by 2.9 percentage points between 2003 and 2011 ($P=.69$) and then narrowed by 6.2 percentage points between 2011 and 2018 ($P=.50$).

Figure 2. Electronic health use by education level, 2003 to 2018. The sample for provider messaging includes 31,672 total responses, and the sample for looking for health information on the Web includes 27,860 total responses. Survey weights were used to generate means reflective of the US population. Bars represent 95% CIs generated using jackknife SEs. Brackets represent the difference in prevalence between the highest and lowest education groups in the first and last years of the analysis. Of the overall respondents, 3% were not included in this analysis because they were missing information on education.

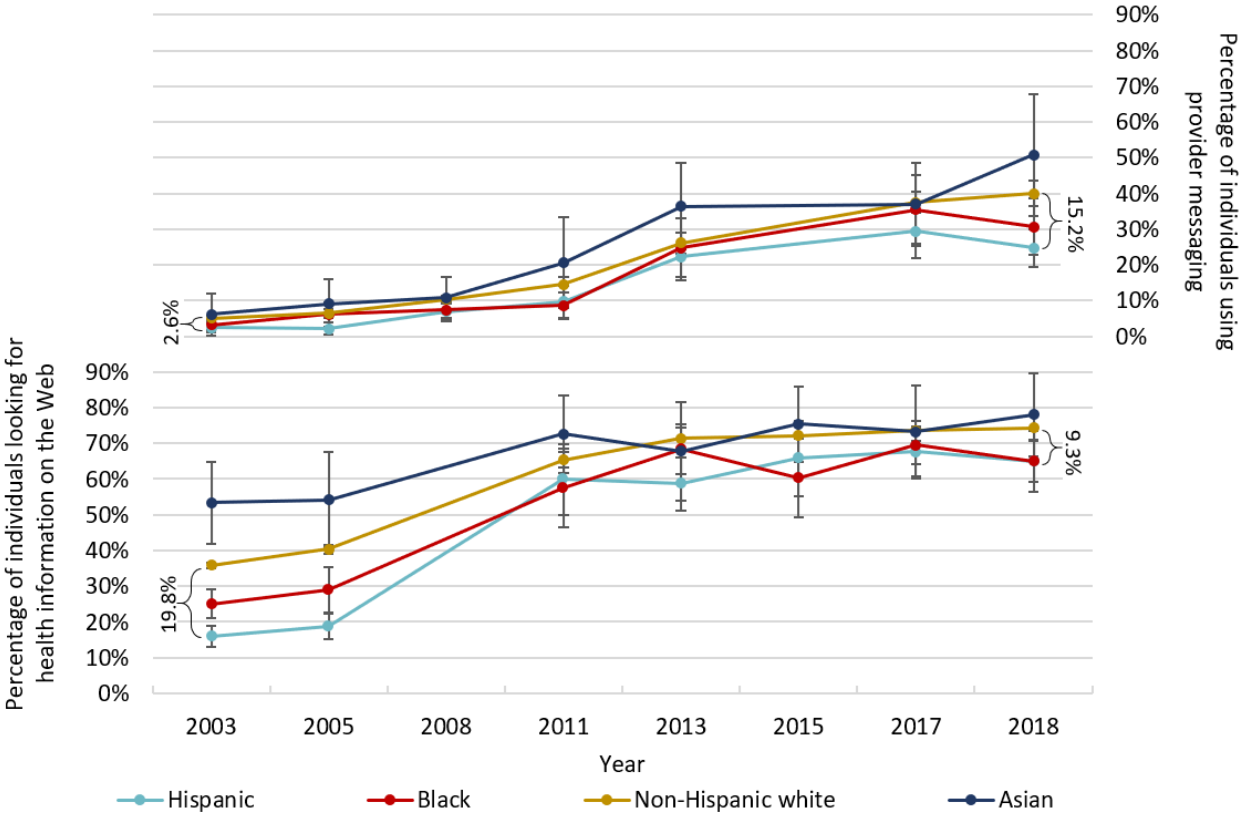


Race and Ethnicity

Use of provider messaging increased across all racial and ethnic groups between 2003 and 2018 (Figure 3, top panel), although the growth was slower among traditionally underserved groups, as was the case with education groups. Among Hispanics, who reported messaging providers least often, the reported rates of provider messaging increased by 22 percentage points, the smallest increase of any group. Meanwhile, among non-Hispanic whites, the reported rates increased by 35 percentage points. As with education, the majority of these increases occurred in the postincentive period: between 2011 and 2018, rates of provider messaging increased by 15 percentage points among Hispanics and by 25 percentage points among non-Hispanic whites. The gap in use of provider messaging between Hispanics and non-Hispanic whites increased by only 2.1 points between 2003 and 2011 (P=.42) but widened an additional 11 points between 2011 and 2018 (P=.01). Trends for blacks closely followed those for Hispanics, whereas trends for Asians more closely resembled those for non-Hispanic whites.

The rates of looking for health information on the Web also increased across all racial and ethnic groups. However, unlike provider messaging, the reported rates of looking for health information on the Web increased most quickly among traditionally underserved racial groups. Among Hispanics, the reported rates increased by 49 percentage points between 2003 and 2018, the greatest increase of any group. Among non-Hispanic whites, the rates increased by 38 percentage points. Paralleling trends across education levels, the majority of this increase occurred in the preincentive period, with rates increasing by 44 percentage points for Hispanics and 30 percentage points for non-Hispanic whites between 2003 and 2011. Therefore, the gap between Hispanic and non-Hispanic whites narrowed by 15 points between 2003 and 2011 (P=.008) and then widened by 3.8 points between 2011 and 2018 (P=.53).

Figure 3. Electronic health use by race and ethnicity. The sample for provider messaging includes 29,484 total responses, and the sample for looking for health information on the Web includes 25,638 total responses. Survey weights were used to generate means reflective of the US population. Bars represent 95% CIs generated using SEs. Brackets represent the difference in prevalence between Hispanic and non-Hispanic white respondents in the first and last years of the analysis. Other and multiracial categories were excluded from this analysis because they did not have at least 100 observations for each sample year. Furthermore, 7% of total respondents were excluded because they did not indicate a race.



Adjusted Associations

Education

The magnitude of the adjusted association between education and provider messaging increased over time but was smaller than the unadjusted differences presented above (Table 1). In 2003 to 2005, college graduates were 6.2 percentage points more likely to use provider messaging than those with less than a high school education (P<.001), whereas in 2017 to 2018, college graduates were 24 percentage points more likely to use provider messaging than those with less than a high school education (P<.001). In comparison, the adjusted association between education and looking for health information on the Web grew less strong over time. In 2003 to 2005, college graduates were 34 percentage points more likely to look for health information on the Web than those with less than a high school education (P<.001). By 2017 to 2018, that difference decreased to 28 percentage points (P<.001).

Race and Ethnicity

As with education, the associations between race and ethnicity and provider messaging increased over time, whereas the associations with looking up health information on the Web lessened. In 2003 to 2005, non-Hispanic whites were 2.0 percentage points more likely to use provider messaging than Hispanics (P=.07), compared with 5.4 percentage points in 2017 to 2018 (P=.05). In contrast, in 2003 to 2005, non-Hispanic whites were 11 percentage points more likely than Hispanics to look for health information on the Web (P<.001). By 2017 to 2018, the difference decreased to 5.2 percentage points (P=.06).

Explanatory Factors

In 2017 to 2018, internet adoption and access to health care providers were strongly associated with both provider messaging and looking for health information on the Web. Physician use of an EMR was associated with a 20 percentage point increase in the likelihood of provider messaging (P<.001) but was not associated with looking up health information on the Web. Including these variables in the model attenuated, but did not eliminate, the associations between education and provider messaging or looking for health information on the Web. Disparities across racial and ethnic groups in provider messaging and looking for health information on the Web were no longer statistically significant when accounting for these explanatory factors.

Discussion

Principal Findings

In nationally representative data from 2003 to 2018, both provider messaging and looking up health information on the Web became more common. However, the *digital divide* in the use of financially incentivized, provider-focused eHealth (provider messaging) widened, whereas the divide in eHealth that is independent of providers and policy-based incentives (looking up health information on the Web) stayed the same across education levels and narrowed across racial and ethnic groups. For all groups, the rates of provider messaging grew more rapidly in the years following the introduction of federal financial incentives, whereas the rates of looking up information on the Web plateaued. Disparities that persisted in 2017 to 2018 were only partially explained by differences in internet adoption, health care access, or provider use of an EMR. These findings indicate that federal incentives may have accelerated growth in the technologies they targeted across all groups, but they may have disproportionately impacted growth among white, well-educated, and wealthier individuals.

Existing theories and models of patient adoption of health technologies suggest that persistent disparities in health technologies stem from systematic differences at the patient, provider, and system levels [22-24]. Financial incentives directed at providers address provider participation in secure messaging, but they may have failed to address several other provider-level factors that differentially impact groups of patients. Although our data included a measure of whether providers maintained an EMR, they did not include measures of whether providers offered secure messaging to patients. Providers serving underresourced communities may not offer secure messaging to patients because they are too busy, are not comfortable with eHealth tools themselves, hold beliefs or biases that these groups of patients are unlikely to use or benefit from eHealth tools, or are less able or willing to change workflows to facilitate the use of sometimes cumbersome messaging tools [25-28]. One key issue is that EMRs vary in quality, and adopting poorer-quality EMRs may make messaging time consuming, difficult, and poorly integrated into existing clinical workflows [29-31]. Public policies, including MU, may have exacerbated the differences in EMR quality. High-resource practices, which often serve wealthier patient groups, were more likely to use advance EMRs, attest to MU, and receive payments [32-35]. In contrast, lower-resource practices may have either viewed the MU criteria as too challenging or have adopted systems that were just good enough to facilitate MU attestation. The challenges of working with poorer-quality EMRs are likely made more difficult by the fact that practices serving low-resource patients tend to have especially limited time for each patient visit [36].

Beyond these important provider-level considerations, evidence indicates that several differences in patient-level factors, such as health literacy, eHealth literacy, attitudes toward Web-based health information, and social norms, also contribute to continued disparities in eHealth use [4,5,10,12,13]. Some related mechanisms may be unique to the growing disparities in the use of provider messaging. For example, evidence suggests that there are barriers to high-quality interpersonal communication between patients who are racial and/or ethnic minorities, have lower levels of education, or have lower incomes and their providers [37-39]. Poorer interpersonal communication may discourage patients from communicating with physicians outside of the clinic setting through eHealth. Finally, although the spread of the internet has narrowed the digital divide [11], African Americans, Hispanics, and low-income individuals are more likely than white or wealthier Americans to rely on smartphones for their internet connections, leading to unreliable access, especially for those who reach their maximum monthly data allotments or have to cancel or suspend phone services because of financial hardship [40-42]. The measure of internet adoption used in this research did not capture these kinds of disparities in access that may impede patients' ability to engage in health technologies. Underresourced patients are also more likely to share devices or use publicly available devices, perhaps sparking concerns about the privacy and security of sharing their health information over the internet. Indeed, a previous study has shown that African Americans are more concerned about health information security than whites and that these concerns predict engagement in eHealth more strongly for African Americans' than they do for whites [13].

As supplements to newer policy efforts under the Merit-Based Incentive Payment System, public policy should consider ways to overcome barriers that are particularly likely to impede provider messaging with patients in lower-resource settings. These might include further development of team-based approaches that move some of the burden of messaging from physicians to other professionals, provider education around strategies to maximize benefit from messaging, and support for provider outreach programs specifically targeting disadvantaged patients. Policy makers may also need to complement provider-facing initiatives with other programs designed to increase digital inclusion. For example, a previous study has shown that lacking internet access in one's neighborhood is a major factor associated with patient portal use [43], suggesting a need for further efforts aimed at increasing reliable, secure access to the internet. Other efforts could target patient-level barriers to adoption, including general skills training to improve proficiency of internet use and targeted training to orient patients to portal use and/or secure messaging. Policy makers might also consider how to provide incentives for smaller EMR vendors to enhance the usability of patient messaging platforms to close the *advanced use* gap between providers in high- versus low-resource settings [17]. Although our data are from the United States, other countries may similarly find that without careful design, the benefits of information technology investment and financial incentives disproportionately benefit some groups over others.

We presented both unadjusted (Figures 1-3) and adjusted (Table 1) associations between demographic variables and eHealth use because of the closely intertwined nature of race, education, and income. Multivariate modeling approaches that attempt to isolate the independent effects of each may obscure important relationships by controlling for parts of the causal pathway. In these data, lower rates of eHealth use among some racial and

ethnic minority groups are mostly accounted for by the inclusion of education, income, and other predictors in adjusted models. This suggests that policies addressing differences by education and income, for example, by increasing eHealth literacy or incentivizing providers' use of secure messaging in low-resource settings, may be most impactful while also decreasing racial disparities.

We report absolute differences in eHealth use between groups, although some previous literature [1,12] has focused on relative differences. Given the rapidly changing rates of use across all groups during this period, absolute differences are more readily interpretable. Still, both absolute and relative differences indicate that the divide in provider messaging is worse relative to the divide in the use of the internet to look up health information. We find that the digital divide in looking for health information on the Web has stayed constant, whereas the divide in provider messaging has grown. In relative terms, the divide in looking up health information on the Web has decreased (eg, from eight-fold in 2005 to two-fold in 2018 across education groups), whereas the divide in provider messaging has stayed fairly constant (eg, from five-fold to four-fold).

Limitations

Our study is subject to a number of limitations. HINTS is a cross-sectional survey that does not allow for longitudinal analysis of change in individuals over time but rather changes by group characteristics; therefore, selection bias in specific survey years may influence our findings. Furthermore, similar to all survey data, HINTS data may be subject to nonresponse or self-reporting bias. Our analysis does not support causal inference: although we highlight that changing rates of eHealth use parallel enactment of public policy related to provider messaging, specifically federal financial incentives to providers, we cannot definitively state that the incentives caused these changes. Similarly, as we have discussed, it is likely that unobserved mediating variables are more proximal causes of the digital divide than the demographic variables measured here. Finally, our analysis was constrained to only 2 eHealth activities that were asked in most years of HINTS. These activities are representative of provider-focused, incentivized eHealth and of eHealth that is independent of providers and not directly incentivized [44], but they differ in other ways as well. For example, looking up health information on the Web was more common in the baseline year than provider messaging, which may in part explain the tapering increases in use in later years. Using additional measures of eHealth activities would bolster our inference that the relationship with federal financial incentives influenced diverging rates of use of each activity, but those measures are not available.

Conclusions

Using recent, nationally representative data on individuals' use of Web-based tools, we identified a growing digital divide in the rate of messaging with health care providers relative to looking for health information on the Web. This indicates that although federal financial incentive initiatives were successful at increasing patient-provider messaging across all groups, they may have also disproportionately benefited socioeconomically advantaged groups. Moving forward, policy makers should consider how redesigned policy initiatives and new policies might reduce disparities in the use of tools intended to facilitate communication between patients and providers.

Acknowledgments

The research reported in this publication was supported by the Agency for Healthcare Research and Quality (AHRQ) and the Patient-Centered Outcomes Research Institute (PCORI) under K12 HS026395 and T32 HS026122. The content is solely the responsibility of the authors and does not necessarily represent the official views of AHRQ or PCORI.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Electronic health use by income, 2003-2018. The sample for provider messaging includes 28,238 total responses, and the sample for looking for health information online includes 25,080 total responses. Survey weights were used to generate means reflective of the US population. Bars represent 95% CIs, generated using jackknife SEs. Income was coded differently in 2003, the lowest group was less than US $25k and the highest was above US $75k, these groups were therefore excluded from the figure in that survey year.

[PDF File (Adobe PDF File), 88 KB - jmir_v21i10e14976_app1.pdf]

Multimedia Appendix 2

Results of multivariable logistic regression models predicting provider messaging and looking for health information on the Web in 3 separate periods: before major public investment (2003-2005), the first years of public support (2011-2013), and recent years (2017-2018). Furthermore, 2 additional models included 3 variables to test potential explanations for remaining disparities in electronic health use in 2017 to 2018: whether respondents were internet users, whether they had seen a doctor in the past year, and whether their health care provider maintained electronic medical records. Analyses were adjusted for individuals' income, age, gender, marital status, insurance status, and health status. Linear probability models were generated using complete case

analyses. Survey weights were used to generate means reflective of the US population. P values were created using jackknife SEs. SEs are presented in parentheses.
[PDF File (Adobe PDF File), 253 KB - jmir_v21i10e14976_app2.pdf]

References

1. Hong YA, Cho J. Has the digital health divide widened? Trends of health-related internet use among older adults from 2003 to 2011. J Gerontol B Psychol Sci Soc Sci 2017 Sep 1;72(5):856-863. [doi: 10.1093/geronb/gbw100] [Medline: 27558403]
2. Prestin A, Vieux SN, Chou WS. Is online health activity alive and well or flatlining? Findings from 10 years of the health information national trends survey. J Health Commun 2015;20(7):790-798. [doi: 10.1080/10810730.2015.1018590] [Medline: 26042588]
3. Powell JA, Darvell M, Gray JA. The doctor, the patient and the world-wide web: how the internet is changing healthcare. J R Soc Med 2003 Feb;96(2):74-76 [FREE Full text] [doi: 10.1258/jrsm.96.2.74] [Medline: 12562977]
4. Levy H, Janke AT, Langa KM. Health literacy and the digital divide among older Americans. J Gen Intern Med 2015 Mar;30(3):284-289 [FREE Full text] [doi: 10.1007/s11606-014-3069-5] [Medline: 25387437]
5. Nguyen A, Mosadeghi S, Almario CV. Persistent digital divide in access to and use of the internet as a resource for health information: results from a California population-based study. Int J Med Inform 2017 Jul;103:49-54. [doi: 10.1016/j.ijmedinf.2017.04.008] [Medline: 28551001]
6. Collins SA, Yoon S, Rockoff ML, Nocenti D, Bakken S. Digital divide and information needs for improving family support among the poor and underserved. Health Informatics J 2016 Mar;22(1):67-77 [FREE Full text] [doi: 10.1177/1460458214536065] [Medline: 24935213]
7. Anthony DL, Campos-Castillo C, Lim PS. Who isn't using patient portals and why? Evidence and implications from a national sample of US adults. Health Aff (Millwood) 2018 Dec;37(12):1948-1954. [doi: 10.1377/hlthaff.2018.05117] [Medline: 30633673]
8. Huerta TR, Walker DM, Johnson T, Ford EW. A time series analysis of cancer-related information seeking: hints from the health information national trends survey (HINTS) 2003-2014. J Health Commun 2016 Sep;21(9):1031-1038. [doi: 10.1080/10810730.2016.1204381] [Medline: 27565190]
9. Greenberg AJ, Serrano KJ, Thai CL, Blake KD, Moser RP, Hesse BW, et al. Public use of electronic personal health information: measuring progress of the Healthy People 2020 objectives. Health Policy Technol 2017 Mar;6(1):33-39 [FREE Full text] [doi: 10.1016/j.hlpt.2016.08.003] [Medline: 28480160]
10. Chesser A, Burke A, Reyes J, Rohrberg T. Navigating the digital divide: a systematic review of ehealth literacy in underserved populations in the United States. Inform Health Soc Care 2016;41(1):1-19. [doi: 10.3109/17538157.2014.948171] [Medline: 25710808]
11. Pew Research Center. 2019. Internet/Broadband Fact Sheet URL: http://www.pewinternet.org/fact-sheet/internet-broadband/ [accessed 2019-08-16]
12. Graetz I, Gordon N, Fung V, Hamity C, Reed ME. The digital divide and patient portals: internet access explained differences in patient portal use for secure messaging by age, race, and income. Med Care 2016 Aug;54(8):772-779. [doi: 10.1097/MLR.0000000000000560] [Medline: 27314262]
13. Senft N, Abrams J, Katz A, Barnes C, Charbonneau DH, Beebe-Dimmer JL, et al. eHealth activity among African American and white cancer survivors: a new application of theory. Health Commun 2019 Jan 2;1(6):1-6. [doi: 10.1080/10410236.2018.1563031]
14. Tarver WL, Menser T, Hesse BW, Johnson TJ, Beckjord E, Ford EW, et al. Growth dynamics of patient-provider internet communication: trend analysis using the health information national trends survey (2003 to 2013). J Med Internet Res 2018 Mar 29;20(3):e109 [FREE Full text] [doi: 10.2196/jmir.7851] [Medline: 29599107]
15. Sudore RL, Mehta KM, Simonsick EM, Harris TB, Newman AB, Satterfield S, et al. Limited literacy in older people and disparities in health and healthcare access. J Am Geriatr Soc 2006 May;54(5):770-776. [doi: 10.1111/j.1532-5415.2006.00691.x] [Medline: 16696742]
16. Heisey-Grove D, King JA. Physician and practice-level drivers and disparities around meaningful use progress. Health Serv Res 2017 Feb;52(1):244-267 [FREE Full text] [doi: 10.1111/1475-6773.12481] [Medline: 26990114]
17. Adler-Milstein J, Holmgren AJ, Kralovec P, Worzala C, Searcy T, Patel V. Electronic health record adoption in US hospitals: the emergence of a digital 'advanced use' divide. J Am Med Inform Assoc 2017 Nov 1;24(6):1142-1148. [doi: 10.1093/jamia/ocx080] [Medline: 29016973]
18. Sandefer RH, Marc DT, Kleeberg P. Meaningful use attestations among us hospitals: the growing rural-urban divide. Perspect Health Inf Manag 2015;12:1f [FREE Full text] [Medline: 26755900]
19. Wieringa S, Greenhalgh T. 10 years of mindlines: a systematic review and commentary. Implement Sci 2015 Apr 9;10:45 [FREE Full text] [doi: 10.1186/s13012-015-0229-x] [Medline: 25890280]
20. Ohno-Machado L. Understanding and mitigating the digital divide in health care. J Am Med Inform Assoc 2017 Sep;24(5):881. [doi: 10.1093/jamia/ocx082]

21. Health Information National Trends Survey | HINTS. Survey Instruments URL: https://hints.cancer.gov/data/survey-instruments.aspx [accessed 2019-08-20]
22. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. MIS Q 2003;27(3):425-478. [doi: 10.2307/30036540]
23. McGinn CA, Grenier S, Duplantie J, Shaw N, Sicotte C, Mathieu L, et al. Comparison of user groups' perspectives of barriers and facilitators to implementing electronic health records: a systematic review. BMC Med 2011 Apr 28;9:46 [FREE Full text] [doi: 10.1186/1741-7015-9-46] [Medline: 21524315]
24. Palacholla RS, Fischer N, Coleman A, Agboola S, Kirley K, Felsted J, et al. Provider- and patient-related barriers to and facilitators of digital health technology adoption for hypertension management: scoping review. JMIR Cardio 2019 Mar 26;3(1):e11951. [doi: 10.2196/11951]
25. Miller Jr DP, Latulipe C, Melius KA, Quandt SA, Arcury TA. Primary care providers' views of patient portals: interview study of perceived benefits and consequences. J Med Internet Res 2016 Jan 15;18(1):e8 [FREE Full text] [doi: 10.2196/jmir.4953] [Medline: 26772771]
26. Wolcott V, Agarwal R, Nelson DA. Is provider secure messaging associated with patient messaging behavior? Evidence from the US army. J Med Internet Res 2017 Apr 6;19(4):e103 [FREE Full text] [doi: 10.2196/jmir.6804] [Medline: 28385681]
27. Heyworth L, Clark J, Marcello TB, Paquin AM, Stewart M, Archambeault C, et al. Aligning medication reconciliation and secure messaging: qualitative study of primary care providers' perspectives. J Med Internet Res 2013 Dec 2;15(12):e264 [FREE Full text] [doi: 10.2196/jmir.2793] [Medline: 24297865]
28. Thompson MJ, Reilly JD, Valdez RS. Work system barriers to patient, provider, and caregiver use of personal health records: a systematic review. Appl Ergon 2016 May;54:218-242. [doi: 10.1016/j.apergo.2015.10.010] [Medline: 26851482]
29. Carayon P, Wetterneck TB, Alyousef B, Brown RL, Cartmill RS, McGuire K, et al. Impact of electronic health record technology on the work and workflow of physicians in the intensive care unit. Int J Med Inform 2015 Aug;84(8):578-594 [FREE Full text] [doi: 10.1016/j.ijmedinf.2015.04.002] [Medline: 25910685]
30. Howard J, Clark EC, Friedman A, Crosson JC, Pellerano M, Crabtree BF, et al. Electronic health record impact on work burden in small, unaffiliated, community-based primary care practices. J Gen Intern Med 2013 Jan;28(1):107-113 [FREE Full text] [doi: 10.1007/s11606-012-2192-4] [Medline: 22926633]
31. Friedberg MW, Chen PG, van Busum KR, Aunon F, Pham C, Caloyeras J, et al. Factors affecting physician professional satisfaction and their implications for patient care, health systems, and health policy. Rand Health Q 2014;3(4):1 [FREE Full text] [Medline: 28083306]
32. Green LA, Potworowski G, Day A, May-Gentile R, Vibbert D, Maki B, et al. Sustaining 'meaningful use' of health information technology in low-resource practices. Ann Fam Med 2015;13(1):17-22 [FREE Full text] [doi: 10.1370/afm.1740] [Medline: 25583887]
33. Samuel CA. Area-level factors associated with electronic health record adoption and meaningful use in the Regional Extension Center Program. J Am Med Inform Assoc 2014;21(6):976-983 [FREE Full text] [doi: 10.1136/amiajnl-2013-002347] [Medline: 24798687]
34. Bishop TF, Press MJ, Mendelsohn JL, Casalino LP. Electronic communication improves access, but barriers to its widespread adoption remain. Health Aff (Millwood) 2013 Aug;32(8):1361-1367 [FREE Full text] [doi: 10.1377/hlthaff.2012.1151] [Medline: 23918479]
35. Kruse CS, Kristof C, Jones B, Mitchell E, Martinez A. Barriers to electronic health record adoption: a systematic literature review. J Med Syst 2016 Dec;40(12):252 [FREE Full text] [doi: 10.1007/s10916-016-0628-9] [Medline: 27714560]
36. Reschovsky JD, O'Malley AS. Do primary care physicians treating minority patients report problems delivering high-quality care? Health Aff (Millwood) 2008;27(3):w222-w231. [doi: 10.1377/hlthaff.27.3.w222] [Medline: 18430747]
37. Penner LA, Eggly S, Griggs JJ, Underwood III W, Orom H, Albrecht TL. Life-threatening disparities: the treatment of black and white cancer patients. J Soc Issues 2012 Jun 25;68(2):328-357 [FREE Full text] [doi: 10.1111/j.1540-4560.2012.01751.x] [Medline: 24319297]
38. Eggly S, Hamel LM, Foster TS, Albrecht TL, Chapman R, Harper FW, et al. Randomized trial of a question prompt list to increase patient active participation during interactions with black patients and their oncologists. Patient Educ Couns 2017 May;100(5):818-826 [FREE Full text] [doi: 10.1016/j.pec.2016.12.026] [Medline: 28073615]
39. Johnson RL, Roter D, Powe NR, Cooper LA. Patient race/ethnicity and quality of patient-physician communication during medical visits. Am J Public Health 2004 Dec;94(12):2084-2090. [doi: 10.2105/ajph.94.12.2084] [Medline: 15569958]
40. Anderson M, Kumar M. Pew Research Center. 2019. Digital Divide Persists Even as Lower-Income Americans Make Gains in Tech Adoption URL: https://www.pewresearch.org/fact-tank/2019/05/07/digital-divide-persists-even-as-lower-income-americans-make-gains-in-tech-adoption/ [accessed 2019-08-19]
41. Perrin A, Turner E. Pew Research Center. 2019. Smartphones Help Blacks, Hispanics Bridge Some – But Not All – Digital Gaps With Whites URL: https://www.pewresearch.org/fact-tank/2017/08/31/smartphones-help-blacks-hispanics-bridge-some-but-not-all-digital-gaps-with-whites/ [accessed 2019-08-21]
42. Smith A. Pew Research Center. 2015. US Smartphone Use in 2015 URL: https://www.pewinternet.org/2015/04/01/us-smartphone-use-in-2015/ [accessed 2018-08-19]

43. Perzynski AT, Roach MJ, Shick S, Callahan B, Gunzler D, Cebul R, et al. Patient portals and broadband internet inequality. J Am Med Inform Assoc 2017 Sep 1;24(5):927-932 [FREE Full text] [doi: 10.1093/jamia/ocx020] [Medline: 28371853]
44. Senft N, Everson J. eHealth engagement as a response to negative healthcare experiences: cross-sectional survey analysis. J Med Internet Res 2018 Dec 5;20(12):e11034 [FREE Full text] [doi: 10.2196/11034] [Medline: 30518513]

Abbreviations

AHRQ: Agency for Healthcare Research and Quality
eHealth: electronic health
EMR: electronic medical record
HINTS: Health Information National Trends Survey
MU: meaningful use
PCORI: Patient-Centered Outcomes Research Institute

Edited by G Eysenbach; submitted 08.06.19; peer-reviewed by A Sheon, T Muto; comments to author 30.07.19; revised version received 13.08.19; accepted 13.08.19; published 07.10.19.

Please cite as:
Senft N, Butler E, Everson J
Growing Disparities in Patient-Provider Messaging: Trend Analysis Before and After Supportive Policy
J Med Internet Res 2019;21(10):e14976
URL: https://www.jmir.org/2019/10/e14976
doi:10.2196/14976
PMID:31593539

Original Paper

Psychosocial Factors Affecting Artificial Intelligence Adoption in Health Care in China: Cross-Sectional Study

Tiantian Ye[1,2], MD; Jiaolong Xue[3,4], PhD; Mingguang He[5,6], MPH, MD, PhD; Jing Gu[7], PhD; Haotian Lin[5], MD, PhD; Bin Xu[8], MSc; Yu Cheng[1,9], PhD

[1]Department of Anthropology, School of Sociology and Anthropology, Sun Yat-sen University, Guangzhou, China

[2]Department of Preventive Ophthalmology, Zhongshan Ophthalmic Center, Sun Yat-sen University, Guangzhou, China

[3]Business School, Sun Yat-sen University, Guangzhou, China

[4]School of Management, Guangdong Ocean University, Zhanjiang, China

[5]State Key Laboratory of Ophthalmology, Zhongshan Ophthalmic Center, Sun Yat-sen University, Guangzhou, China

[6]Center for Eye Research Australia, University of Melbourne, Melbourne, Australia

[7]Department of Epidemiology and Medical Statistics, School of Public Health, Sun Yat-sen University, Guangzhou, China

[8]Guangzhou Yuexiu Center for Disease Control and Prevention, Guangzhou, China

[9]Department of Medical Humanities, The Seventh Affiliated Hospital, Sun Yat-sen University, Shenzhen, China

Corresponding Author:
Yu Cheng, PhD
Department of Medical Humanities
The Seventh Affiliated Hospital
Sun Yat-sen University
No 628, Zhenyuan Rd Guangming (New) Dist
Shenzhen, 518107
China
Phone: 86 02084114275
Email: chengyu@mail.sysu.edu.cn

Abstract

Background: Poor quality primary health care is a major issue in China, particularly in blindness prevention. Artificial intelligence (AI) could provide early screening and accurate auxiliary diagnosis to improve primary care services and reduce unnecessary referrals, but the application of AI in medical settings is still an emerging field.

Objective: This study aimed to investigate the general public's acceptance of ophthalmic AI devices, with reference to those already used in China, and the interrelated influencing factors that shape people's intention to use these devices.

Methods: We proposed a model of ophthalmic AI acceptance based on technology acceptance theories and variables from other health care–related studies. The model was verified via a 32-item questionnaire with 7-point Likert scales completed by 474 respondents (nationally random sampled). Structural equation modeling was used to evaluate item and construct reliability and validity via a confirmatory factor analysis, and the model's path effects, significance, goodness of fit, and mediation and moderation effects were analyzed.

Results: Standardized factor loadings of items were between 0.583 and 0.876. Composite reliability of 9 constructs ranged from 0.673 to 0.841. The discriminant validity of all constructs met the Fornell and Larcker criteria. Model fit indicators such as standardized root mean square residual (0.057), comparative fit index (0.915), and root mean squared error of approximation (0.049) demonstrated good fit. Intention to use (R^2=0.515) is significantly affected by subjective norms (beta=.408; *P*<.001), perceived usefulness (beta=.336; *P*=.03), and resistance bias (beta=–.237; *P*=.02). Subjective norms and perceived behavior control had an indirect impact on intention to use through perceived usefulness and perceived ease of use. Eye health consciousness had an indirect positive effect on intention to use through perceived usefulness. Trust had a significant moderation effect (beta=–.095; *P*=.049) on the effect path of perceived usefulness to intention to use.

Conclusions: The item, construct, and model indicators indicate reliable interpretation power and help explain the levels of public acceptance of ophthalmic AI devices in China. The influence of subjective norms can be linked to Confucian culture, collectivism, authoritarianism, and conformity mentality in China. Overall, the use of AI in diagnostics and clinical laboratory analysis is underdeveloped, and the Chinese public are generally mistrustful of medical staff and the Chinese medical system.

Stakeholders such as doctors and AI suppliers should therefore avoid making misleading or over-exaggerated claims in the promotion of AI health care products.

(J Med Internet Res 2019;21(10):e14316) doi:10.2196/14316

KEYWORDS

artificial intelligence; adoption; technology acceptance model; structural equation model; intention; subjective norms; trust; moderation

Introduction

Background

As part of the fourth industrial revolution, artificial intelligence (AI) has achieved massive progress and explosive growth. It is actively applied in health care to perform a wide range of functions such as patient administration and monitoring, clinical decision support, risk prediction, medical error reduction, health care intervention, and productivity improvement [1,2]. These potential benefits could contribute greatly to primary care services in China, where the health system is facing great challenges owing to an aging population and an increase in chronic noncommunicable diseases [3].

This challenge is especially crucial for eye health management in China, where rates of blindness and vision impairment are the highest in the world and age-related eye diseases such as cataracts, diabetic retinopathy (DR), and juvenile myopia are increasingly common [4]. Most of these diseases cannot be diagnosed in primary care institutions, so patients seek direct care from ophthalmologists in tertiary hospitals without a referral. Data from 1 survey in Shanghai showed that on average there are only 0.09 ophthalmologists and 0.1 primary eye care (PEC) providers for every 10,000 people [5]. Of the available ophthalmologists and PEC providers, 82.9% majored in public health, nursing, or internal medicine and have not had specialist ophthalmic training. The situation is even worse in areas of western China, such as Tibet and Inner Mongolia, where the high prevalence of blindness and poor vision has become a serious public health issue. It is vital to establish and maintain an appropriate, effective eye care program in these areas [6,7].

Researchers have demonstrated that the performance of image-based AI devices can reach or even surpass that of experts [8-10]. The number of effective programs and policies to prevent blindness in China has increased [11], and a number of ophthalmic AI devices are available in clinical scenarios, such as EyeGrader (Center for Eye Research Australia, Melbourne, Australia) for the detection of DR [12] and CC-Cruiser (Zhongshan Ophthalmic Center, Guangzhou, China) for congenital cataracts [13]. Stakeholders such as doctors and AI suppliers are trying to apply these devices in clinical settings such as health check centers, community health centers (CHCs), schools, optical stores, and grassroots hospitals in rural China [14]. As no prior studies have been conducted on the implementation of ophthalmic AI devices in the Chinese context, we have briefly described the results of our formative qualitative studies of 3 CHCs where an ophthalmic AI device was used (unpublished). During the implementation period from April 1 to December 31, 2018, the total number of people who signed the Service of Community Family Physician was 63,034. We found that the low number of patients who chose to receive AI screening (3067 out of 63,034) could reflect public unwillingness to use these devices, though AI screening was not systematically offered by physicians. In interviews with patients, we found that patients were unwilling to undergo this process unless it was provided free of charge, arranged by their work unit, or they could attend the screening in a group with other people.

Above all, in recent years, AI health care researchers have focused on technical innovation and clinical results, without considering the human context or ethical challenges that are invariably involved in any complex health care system. Many real-world issues need to be assessed in the implementation phase, most notably the extent to which patients or the public accept AI and the challenges involved in protecting patient privacy and confidential medical information. Thus, understanding the factors that influence public acceptance of (or resistance to) AI devices in the Chinese social and cultural context will help government agencies and health care administrators to devise appropriate intervention strategies to minimize user resistance and its negative effects on health care policy.

Objective

The aims of this study were to develop and test a model investigating the factors that drive the public's acceptance of ophthalmic AI devices, with reference to those already used in primary care institutions in China. In particular, we aimed to evaluate how subjective norms, resistance bias, and trust contribute to the relationships among these factors in the Chinese cultural context.

Theoretical Background and Hypothesis Development

Many technology adoption models have been proposed to explain user adoption of new technology and to assess the factors that can affect user acceptance [15]; examples include the Technology Acceptance Model (TAM) [16,17], Theory of Planned Behavior (TPB) [18,19], and the Unified Theory of Acceptance and Use of Technology (UTAUT) [20]. Many medical information researchers have modified and combined models or added new constructs to carry out studies in domains such as telemedicine [21-23], clinical decision systems [24,25], electronic health care records [26-29], mobile medical information systems [30-32], and personal digital assistants [33-35].

Studies of the acceptance of new health care technology have identified influential factors and reliable correlations between those factors and the acceptance or usage of new technology. However, very few studies have been carried out in relation to

AI technology. As ophthalmic AI devices are an emerging technology, this study uses the following theories and constructs to evaluate these influential factors and facilitate the application of AI within primary health care institutions.

Technology Acceptance Model Theories

The TAM is the most widely applied model to describe consumer acceptability of information technology [36]. The original model, developed by Fred D Davis in 1989 [16], revealed that perceived usefulness (PU; defined as the perception that using a system leads to enhanced job performance) and perceived ease of use (PEOU; defined as the perception that using a system will be free of effort) were 2 basic determinants of people's acceptance of new technology, which is now commonly evaluated by behavioral intention to use (IU; defined as an individual's motivation or willingness to exert effort to perform the target behavior) [15,16]. Then on, many researchers have added, modified, or deleted some variables to synthesize new models to fit their studies, such as TAM 2, TAM 3, and UTAUT. However, many researchers found that both PU and PEOU had a direct effect on IU without a mediation effect of attitude, and attitude was deleted in the following TAMs [36,37]. In our study, although the purpose was to understand the Chinese public's acceptance of AI devices, as most regions do not have access to these devices, our final dependent variable was IU, rather than actual usage behavior as indicated in the TAMs.

IU is now commonly used to refer to acceptance and is considered to reliably predict actual use; it is sometimes the only measured outcome of interest in TAM-related studies [15]. Studies have shown that PU and PEOU exert considerable positive influence on IU, and PEOU has an effect on PU [15,16]. We thus proposed the following hypotheses:

- H1: Perceived usefulness positively affects the public's intention to use ophthalmic AI devices.
- H2a: Perceived ease of use positively affects the public's intention to use ophthalmic AI devices.
- H2b: Perceived ease of use positively affects the public's perception of the usefulness of ophthalmic AI devices.

Theory of Planned Behavior

The Theory of Planned Behavior (TPB) states that an individual's behavioral intention (similar to IU in TAMs) is determined by attitude, perceived behavioral control (PBC; the extent to which people have control over engaging in the behavior) and subjective norms (SN; defined as perceptions of whether others think one should engage in a behavior) [38-40]. TPB, a more comprehensive version of the Theory of Reasoned Action (TRA) [41], allows us to examine the influence of personal determinants and social surroundings as well as nonvolitional determinants on IU [42]. As an extension of the TRA, TPB has been one of the most widely tested models of the factors influencing health-related behavior [40]. SN has a direct effect on IU in the UTAUT and TPB models and an indirect impact on IU through PEOU in many integrated models [15,36,43]. PBC has a positive effect on IU in the TPB and UTAUT models [15,44]. However, when combined with TAMs, PBC also has an indirect effect through PEOU [45,46]. Therefore, we proposed the following integrated hypotheses:

- H3a: Subjective norms positively affect the public's intention to use ophthalmic AI devices directly.
- H3b: Subjective norms positively affect the public's perception of the ease of use of ophthalmic AI devices.
- H4a: Perceived behavioral control positively affects the public's intention to use ophthalmic AI devices directly.
- H4b: Perceived behavioral control positively affects the public's perception of the ease of use of ophthalmic AI devices.

Health Belief Model and Eye Health Consciousness

The health belief model (HBM) [47] was initially designed to "understand the widespread failure of people to accept preventives or screening tests for the early detection of asymptomatic disease" [48]. In later studies, the HBM was used to predict more general health-related behaviors, to understand why individuals did or did not engage in these actions, and to explain and predict the acceptance of health and medical care recommendations [36,48,49]. Health consciousness is defined as the "degree to which health concerns are integrated into a person's daily activities and health-conscious people are aware of and concerned about their wellness, resulting in a better motivation to improve or maintain their health" [49]. Health beliefs and concerns have an indirect effect on behavioral intention to use health information technology via the remote mediation effect of perceived health threat (PHT) and PU [36]. One study in China examined patients' acceptance of mobile phone health technology for chronic disease management and showed that PHT had a significant positive effect on PU together with a positive effect directly on IU [31]. In that study, PHT referred to patients' awareness and care of the health condition and its potential consequences. The items in their construct also covered a person's degree of consciousness, beliefs, and awareness of hypertension and health management and asked the participants if they were aware of or concerned about blood pressure and would make efforts to manage hypertension. Therefore, we modified these items to fit the eye care context and defined this construct as eye health consciousness (EHC). We thus proposed the following hypothesis:

- H5a: Eye health consciousness positively affects the public's intention to use ophthalmic AI devices directly.
- H5b: Eye health consciousness positively affects the public's perception of the usefulness of ophthalmic AI devices directly.

Dual Factor Theory and Status Quo Bias Theory

The above health behavior theories focus almost exclusively on users' positive (enabling) perceptions in relation to new technology usage and ignore negative (inhibiting) factors [31,50]. However, in the Dual Factor Theory (DFT), potential users' information technology usage considerations are based on a simultaneous examination of both enabling and inhibiting factors [31]. Inhibitors discourage information systems (IS) usage when present but do not necessarily favor usage when absent. They are not quite the opposite of enablers but are qualitatively distinct constructs that are independent of but may coexist with enablers [51]. Perceived risk (PR) refers to the combination of uncertainty and the seriousness of an outcome in relation to performance, safety, and psychological or social

uncertainties, which have a negative influence on IU and are thus barriers to adoption [28,52,53]. Status Quo Bias (SQB) theory aims to explain people's preference for maintaining their current status or situation and provides a set of useful theoretical explanations for understanding the impact of incumbent system (IS) use as an inhibitor of new IS acceptance. For example, data on the selection of health plans by faculty members reveal that SQB is substantial in important real-world decisions [54], so several studies have modified their models by supplementing the negative (inhibiting) constructs of SQB theory with user resistance factors that are a type of inhibitor [51,55]. The 2 main inhibitors are regret avoidance (lessons from experiences that have taught individuals to avoid regrettable consequences) and inertia (an individual's attachment to his or her current situation even if there are better alternatives or incentives to change) [51,55]. Resistance to change (RTC) refers to people's attempts to maintain their previous behaviors or habits that are connected to their past experiences when facing change [31,56-58]. RTC has been confirmed as a major barrier for electronic health and mobile health adoption [56-58]. We integrated these factors into 1 inhibitor, resistance bias (RB), defined as people's resistance to use a new technology owing to biases such as regret avoidance, inertia, and RTC. We thus proposed the following hypothesis:

- H6: Perceived risk negatively affects the public's intention to use ophthalmic AI devices.
- H7: Resistance bias negatively affects the public's intention to use ophthalmic AI devices.

Trust as a Moderator in the Chinese Social Context

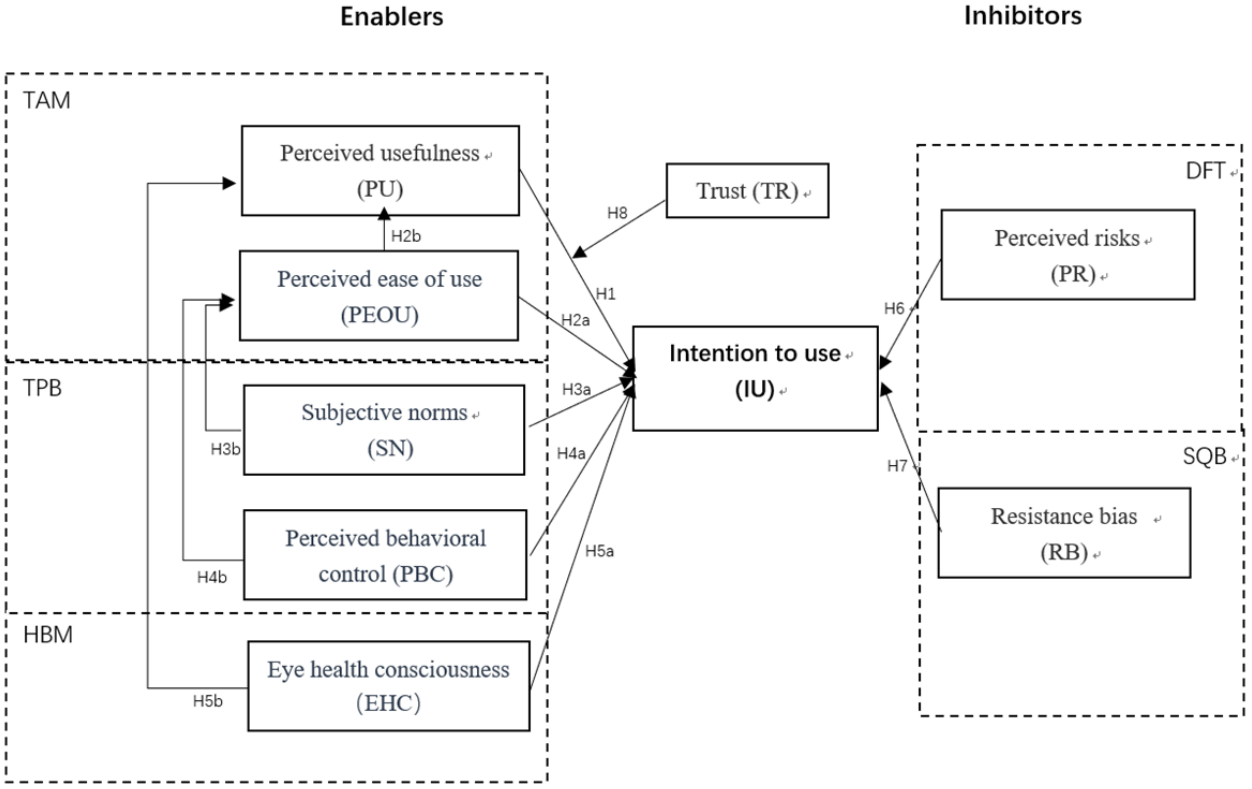
Trust is defined as the belief that someone or something is honest, reliable, good, and effective, or the desire to depend on someone or something for security [44]. Various studies show that it has a direct or indirect mediation effect on user intention or adoption of new technology [52,59-61]. With the increasing proliferation of AI applications in daily life, consideration of trust is essential because it is likely to be a critical factor in the acceptance of consumer products such as home automation, personal robots, and automotive automation [62,63]. Moreover, acceptance behaviors for technologies are controlled and moderated by cultural traits [64,65].

In China, patients' trust of physicians is lower than in Western countries [66] and has become a serious social problem [67,68]. Trust in applied AI is an evolving phenomenon, and cognitive compatibility, trialability, and usability are the main factors related to trust in a technology [63]. The public's trust might play a more complicated role in relation to AI devices in China, affecting the factors that influence IU. In the field of health care research, no previous studies have tested trust of the public as a moderator between PU and IU in China. However, Cuadrado identified the moderating effects of trust, showing that trust levels strengthened the negative effect of prosocialness on selfish irrigation strategies [69]. Although irrigation is unrelated to health or AI, it provides evidence and the possibility that trust might have a potential moderating effect in our context. Thus, we proposed a new hypothesis:

- H8: Trust of physicians moderates the effect of perceived usefulness on the public's intention to use ophthalmic AI devices.

Overall, this study proposes and evaluates 12 hypotheses with TAM and TPB as the underpinning theories (Figure 1). We added the constructs EHC, PR, and RB from HBM, DFT, and SQB, respectively, to fit our context as these constructs have been validated in the previous studies in China or other parts of Asia. Trust was also added as a moderator to reflect the significance of physician-patient relationships in the Chinese context. The selection of variables from relevant theories and the development of our model are shown in Figure 1.

Figure 1. Variables from relevant theories and development of our model for ophthalmic artificial intelligence device acceptance. DFT: Dual Factor Theory; HBM: health belief model; SQB: status quo bias; TAM: Technology Acceptance Model; TPB: Theory of Planned Behavior.



Methods

Participants and Sampling

Potential end users of ophthalmic AI devices in China were recruited if they (1) resided in China (including both urban and rural areas of different provinces and people in all age groups and career types); (2) could read and write in Chinese; (3) had a mobile phone or sufficient internet access; and (4) were not ophthalmic medical staff such as ophthalmologists or nurses.

On the basis of these criteria, we worked with a Web-based company to recruit participants. We calculated the required number of participants based on a sample size rule of thumb for structural equation modeling of 10 times the number of participants as items [70]. As our survey had 32 items, the required number of participants was more than 320. The company distributed the survey to 925 potential participants from January 20 to 24, 2019. The company used simple random sampling of people who were mobile phone users. Its sample database was our source for randomly sampling, which has more than 2.6 million members. Technicians sent selected participants a direct message over a popular messaging platform (WeChat) with the link to the invitation of our questionnaire during certain times on data collection days. Surveys could be completed by potential participants using WeChat. The survey company's website [71] showed the information about the sample source, which was verified and randomized with different job categories. Every day, more than one million people answered questionnaires on this survey platform.

The criteria for determining the completeness of a questionnaire included (1) each account responded only once; (2) the response time was longer than 300 seconds to exclude perfunctory respondents; (3) one *identifying item* randomly selected from an *item bank*, such as *please select the right alphabetical sequence of the following letters: bcdefg*, had to be answered correctly; and (4) anyone choosing *ophthalmic medical staff* in the final *identification item* was excluded.

Measurement

The 9 constructs in the hypothesis model were measured by 32 questionnaire items. Each item measured only 1 construct (variable or factor). All items were sourced from the relevant literature related to consumer technology acceptance research, with some changes to fit the ophthalmic AI context (Table 1). Items in English were translated into Chinese by 1 researcher and checked by 3 other researchers, and 1 researcher then back translated the items into English to check if the original meaning was retained. All researchers are bilingual fluent in English and Chinese. All items were measured on 7-point Likert scales ranging from (1) strongly disagree to (7) strongly agree.

Table 1. Constructs, items, and references of the measurements.

Construct	Definition and items	References
Perceived usefulness (PU)	The degree to which a person believes that the use of ophthalmic AI[a] devices would enhance his or her personal or job performance	[15,16,31,72,73]
PU1	Ophthalmic AI devices would help me to cope with preventable eye diseases at an early stage	[16,31]
PU2	Ophthalmic AI devices would provide detailed information and images of my eyes, which would be very useful for me	[16,31]
PU3	Ophthalmic AI devices would help the medical institutions to recognize more treatable eye patients	[16,31]
PU4	Ophthalmic AI devices would improve primary health care for health departments and save money	[16,31]
PU5	Ophthalmic AI devices would be a good supplement to traditional health care approaches and fit with my medical philosophy	[16,31]
PU6	Ophthalmic AI devices would fit my demand for eye health management	[16,31]
PU7	Ophthalmic AI devices would achieve the same results as face-to-face diagnosis with an ophthalmologist	[16,31]
Perceived ease of use (PEOU)	The degree to which a person believes that ophthalmic AI devices would be easy to use	[15,16,31,72]
PEOU1	I find the instructions for ophthalmic AI devices easy, clear, and understandable	[16,31]
PEOU2	Ophthalmic AI devices would offer a more convenient way for me to cope with my eye disease without queuing for registration in hospitals and would save me time and money	[16,31]
Perceived behavioral control (PBC)	Perception of internal and external resource constraints to using ophthalmic AI devices, or the availability of skills, resources, and opportunities necessary to use them	[15,18,32,44]
PBC1	I have enough knowledge to recognize whether the results of the report are reliable	[15,44]
PBC2	I would receive appropriate technical assistance when encountering any difficulties in using ophthalmic AI devices or understanding the report	[15,44]
PBC3	I would be able to use ophthalmic AI devices independently as long as I had enough time and made an effort to learn	[15,44]
Subjective norms (SN)	Perception of important (or relevant) others' beliefs about my use of ophthalmic AI devices	[15,18,43,44,52]
SN1	People who are important to me (family members, relatives, and close friends) think that I should use ophthalmic AI devices	[15,44]
SN2	My colleagues or peers think that I should use ophthalmic AI devices	[15,44]
SN3	My leaders or superiors think that I should use ophthalmic AI devices	[15,44]
Trust (TR)	The extent to which an individual believes that using ophthalmic AI devices is secure, reliable, effective, and poses no privacy threats	[44,52]
TR1	I would trust that with big data and deep learning, ophthalmic AI devices could deliver a reliable report after analyzing my eye health images	[44,52]
TR2	I would trust that ophthalmic AI devices are more accurate and reliable than human ophthalmologists, because they do not make subjective or empirical errors	[44,52]
TR3	I would trust that stakeholders and reliable third parties would ensure the security and privacy of my personal data, health information, and images	[44,52]
Resistance bias (RB)	Resistance to a new technology owing to biases such as regret avoidance, inertia, and resistance to change	[31,51,56-58]
RB1	I don't want ophthalmic AI devices to change how I deal with eye diseases because I can't be bothered and they are unfamiliar to me	[31,51]
RB2	I don't want to use ophthalmic AI devices because from past experience, these new high-tech products always fall flat during practical applications	[31,51]
RB3	I might regret trying to use these ophthalmic devices because they could waste my time and effort	[31,51]
Eye health consciousness (EHC)	Awareness and care of eye health conditions, and the degree to which eye health concerns are integrated into a person's daily activities	[31,49]
EHC1	I am aware of and very concerned about my eye health	[31,49]
EHC2	I would make efforts to manage my eye health	[31,49]

Construct	Definition and items	References
Perceived risks (PR)	A combination of uncertainty and seriousness of an outcome in relation to performance, safety, psychological or social uncertainties	[28,52,53,73]
PR1	There is a possibility of malfunction and performance failure, so they might fail to deliver accurate diagnoses or recommendations and could increase conflicts between members of the public and medical institutions	[52,53]
PR2	I am concerned that my personal information and health details would be insecure and could be accessed by stakeholders or unauthorized persons, leading to misuse and discrimination	[52,53]
PR3	Considering the difficulties involved in taking high-quality images for AI analysis, I think there is a risk of incorrect screening results	[52,53]
PR4	Given the vision problems I possibly already have, such as visual fatigue, dry eye, or presbyopia, I might find it hard to read the printed or electronic report from ophthalmic AI devices	[52,53]
PR5	Because I might have difficulty understanding the screening report correctly by myself, it might increase my anxiety about my eye health	[52,53]
PR6	Because practitioners with little ophthalmic knowledge might find it difficult to understand the screening report and explain the terminology and results to me, they might increase my anxiety of about my eye health	[52,53]
Intention to use (IU)	An individual's motivation or willingness to exert effort to use ophthalmic AI devices	[15,43,44]
IU1	I intend to use ophthalmic AI devices as my first choice if I feel eye discomfort	[15,44]
IU2	I will encourage my friends/relatives to use ophthalmic AI devices first if they feel eye discomfort	[15,44]
IU3	I will encourage healthy people to use ophthalmic AI devices for eye health path screening	[15,44]

[a]AI: artificial intelligence.

The first page of the questionnaire provided an overview of the study background, purpose, voluntary nature, and anonymity, and asked respondents to indicate their consent. The participants were assured that the questionnaires would only be used by the researchers and would not be accessible to anyone else. On the second page of the questionnaire, we provided a brief introduction to ophthalmic AI devices, including their general functions and operating procedures, with photographs to help instruct the participants. Table 1 shows the constructs and items of the questionnaire and the literature references. We paid ¥12 (US $1.5) to the survey company for each of the 474 completed questionnaires. The company then paid each participant ¥4 (US $0.5). Our Web-based survey was in accordance with the required Checklist for Reporting Results of Internet E-Surveys (Multimedia Appendix 1). Ethical approval was obtained from the Ethics Committee of the Zhongshan Ophthalmic Center, Sun Yat-Sen University.

Data Analysis

SPSS version 25.0 was used to analyze the descriptive statistics. Model evaluation involved a 2-step analysis [74] using Amos 21.0 software by (1) evaluating item and construct reliability and validity via confirmatory factor analysis of the measurement model and (2) evaluating the structural model's path effects, significance, and goodness of fit and mediation and moderation effects.

Results

Demographic Results

We distributed Web-based surveys to 925 potential participants, and 732 individuals participated in the survey (rate of participation, 79.1%, 732/925). Of these, 474 (rate of completion, 64.8%, 474/732) participants who completed the questionnaire and met the criteria were used for the SEM analysis. The participants' demographic characteristics are represented in Table 2. The participants' geographical origins are shown in Table 3.

Table 2. Demographic results.

Characteristics	Values, n (%)
Gender	
Male	169 (35.7)
Female	305 (64.3)
Age (years)	
<18	3 (0.6)
18-25	128 (27.0)
26-30	132 (27.8)
31-40	175 (36.9)
41-50	23 (4.9)
51-60	11 (2.3)
>60	2 (0.4)
Education	
Middle school	4 (0.8)
High school	8 (1.7)
Three-year college	64 (13.5)
Bachelor’s degree	341 (71.9)
Master’s degree	54 (11.4)
Doctoral degree	3(0.6)

Table 3. Geographical origins of participants (N=474).

Province	Value, n (%)
Guangdong	80 (16.9)
Beijing	67 (14.1)
Shanghai	38 (8.0)
Jiangsu	37 (7.8)
Shandong	28 (5.9)
Zhejiang	26 (5.5)
Sichuan	22 (4.6)
Henan	17 (3.6)
Hubei	17 (3.6)
Liaoning	17 (3.6)
Chongqing	16 (3.4)
Anhui	15 (3.2)
Hunan	13 (2.7)
Shaanxi	13 (2.7)
Hebei	10 (2.1)
Fujian	9 (1.9)
Heilongjiang	8 (1.7)
Jiangxi	8 (1.7)
Shanxi	8 (1.7)
Jilin	5 (1.1)
Tianjin	5 (1.1)
Guangxi	4 (0.8)
Yunnan	4 (0.8)
Guizhou	2 (0.4)
Gansu	1 (0.2)
Hainan	1 (0.2)
Inner Mongolia	1 (0.2)
Ningxia	1 (0.2)
Xinjiang	1 (0.2)

The Effect of Education on Intention to Use

The results of a single-factor analysis of variance showed that the main effect of education on IU was not significant ($F_{5,468}$=0.316; P>.05) and that each group of education had no significant difference in terms of IU, with means from 4.750 to 5.204, as shown in Table 4. As predicted, the results of post hoc comparisons revealed that the effect of education on IU was not significant as shown in Table 5.

Table 4. Descriptive statistics of the effect of education on intention to use.

Diploma	Total	Mean (SD)	SE	95% CI for mean	Minimum	Maximum
Middle school	4	4.750 (1.912)	0.956	1.707 to 7.793	3.000	7.000
High school	8	5.375 (1.408)	0.498	4.198 to 6.552	2.333	6.667
Three-year college	64	5.167 (0.914)	0.114	4.938 to 5.395	2.333	7.000
Bachelor’s degree	341	5.199 (1.000)	0.054	5.093 to 5.306	1.000	7.000
Master’s degree	54	5.204 (1.084)	0.148	4.908 to 5.500	2.333	6.667
Doctoral degree	3	4.778 (1.347)	0.778	1.431 to 8.124	3.333	6.000

Table 5. Post hoc multiple comparisons of the effect of education on intention to use (IU; dependent variable: IU Method: Scheffe).

Diploma (I), diploma (J)	Mean difference (I-J)	SE	*P* value	95% CI
Middle school				
High school	–0.625	0.622	.96	–2.704 to 1.454
Three-year college	–0.417	0.524	.99	–2.167 to 1.333
Bachelor's degree	–0.449	0.511	.98	–2.157 to 1.258
Master's degree	–0.454	0.527	.98	–2.213 to 1.306
Doctoral degree	–0.028	0.776	>.99	–2.621 to 2.566
High school				
Middle school	0.625	0.622	.96	–1.454 to 2.704
Three-year college	0.208	0.381	>.99	–1.065 to 1.482
Bachelor's degree	0.176	0.363	>.99	–1.039 to 1.390
Master's degree	0.171	0.385	>.99	–1.115 to 1.458
Doctoral degree	0.597	0.688	.98	–1.702 to 2.896
Three-year college				
Middle school	0.417	0.524	.99	–1.333 to 2.167
High school	–0.208	0.381	>.99	–1.482 to 1.065
Bachelor's degree	–0.033	0.138	>.99	–.495 to .430
Master's degree	–0.037	0.188	>.99	–.664 to .590
Doctoral degree	0.389	0.600	>.99	–1.617 to 2.395
Bachelor's degree				
Middle school	0.449	0.511	.98	–1.258 to 2.157
High school	–0.176	0.363	>.99	–1.390 to 1.039
Three-year college	0.033	0.138	>.99	–.430 to .495
Master's degree	–0.004	0.149	>.99	–.502 to .493
Doctoral degree	0.422	0.589	>.99	–1.547 to 2.391
Master's degree				
Middle school	0.454	0.527	.98	–1.306 to 2.213
High school	–0.171	0.385	>.99	–1.458 to 1.115
Three-year college	0.037	0.188	>.99	–.590 to .664
Bachelor's degree	0.004	0.149	>.99	–.493 to .502
Doctoral degree	0.426	0.603	>.99	–1.588 to 2.440
Doctoral degree				
Middle school	0.028	0.776	>.99	–2.566 to 2.621
High school	–0.597	0.688	.98	–2.896 to 1.702
Three-year college	–0.389	0.600	>.99	–2.395 to 1.617
Bachelor's degree	–0.422	0.589	>.99	–2.391 to 1.547
Master's degree	–0.426	0.603	>.99	–2.440 to 1.588

Measurement Model

Maximum likelihood estimation was used to test the factor loadings, measurement reliability, convergent validity, and discriminant validity. Table 6 presents a summary of the significance tests, item reliability, composite reliability (CR), and convergence validity. The standardized factor loadings of items are between 0.583 and 0.869, with good item reliability. The CR values of the 9 constructs range from 0.673 to 0.841, approaching or exceeding 0.7 [75] All constructs have acceptable internal consistency. Most constructs have an average variance extracted (AVE) value higher than the threshold of 0.5, which confirms the constructs' convergent validity.

Table 6. Descriptive statistics of variables, items, and convergent validity.

Construct, item	Mean	Significant test of parameter estimation				Item reliability		Composite reliability, CR[d]	Convergence validity, AVE[e]
		Unstd[a]	SE	Unstd/SE	*P* value	STD[b]	SMC[c]		
Perceived usefulness (PU)								0.841	0.431
PU1	6.095	1	__[f]	—	—	0.663	0.44		
PU2	6.171	1.076	0.09	11.972	<.001	0.638	0.407		
PU3	6.118	1.118	0.094	11.926	<.001	0.629	0.396		
PU4	5.859	1.344	0.118	11.386	<.001	0.605	0.366		
PU5	5.873	1.419	0.116	12.222	<.001	0.656	0.43		
PU6	5.77	1.518	0.115	13.149	<.001	0.727	0.529		
PU7	5.091	1.958	0.156	12.512	<.001	0.672	0.452		
Perceived ease of use (PEOU)								0.68	0.516
PEOU1	5.715	1	—	—	—	0.685	0.469		
PEOU2	5.762	1.109	0.119	9.313	<.001	0.75	0.562		
Perceived behavioral control (PBC)								0.673	0.408
PBC1	4.62	1	—	—	—	0.71	0.504		
PBC2	5.38	0.748	0.073	10.318	<.001	0.605	0.366		
PBC3	5.015	0.935	0.092	10.178	<.001	0.596	0.355		
Subjective norms (SN)								0.758	0.512
SN1	5.16	1	—	—	—	0.704	0.496		
SN2	5.2	1.122	0.085	13.162	<.001	0.764	0.584		
SN3	5.023	1.009	0.085	11.884	<.001	0.675	0.456		
Trust (TR)								0.691	0.429
TR1	5.359	1	—	—	—	0.583	0.34		
TR2	4.595	1.697	0.166	10.228	<.001	0.732	0.536		
TR3	4.975	1.349	0.141	9.551	<.001	0.642	0.412		
Resistance bias (RB)								0.767	0.524
RB1	2.319	1	—	—	—	0.683	0.466		
RB2	2.479	1.368	0.109	12.567	<.001	0.762	0.581		
RB3	2.259	1.133	0.093	12.123	<.001	0.724	0.524		
Eye health consciousness (EHC)								0.766	0.625
EHC1	6.051	1	—	—	—	0.876	0.767		
EHC2	5.724	0.859	0.136	6.317	<.001	0.694	0.482		
Perceived risks (PR)								0.837	0.461
PR1	3.962	1	—	—	—	0.711	0.506		
PR2	3.979	0.932	0.073	12.814	<.001	0.639	0.408		
PR3	3.804	1.081	0.075	14.468	<.001	0.738	0.545		
PR4	3.308	0.968	0.081	12.025	<.001	0.621	0.386		
PR5	4.217	1.089	0.082	13.34	<.001	0.705	0.497		
PR6	3.544	0.931	0.075	12.483	<.001	0.654	0.428		
Intention to use (IU)								0.753	0.506
IU1	4.977	1	—	—	—	0.743	0.552		
IU2	5.251	0.999	0.069	14.436	<.001	0.769	0.591		

Construct, item	Mean	Significant test of parameter estimation				Item reliability		Composite reliability, CR[d]	Convergence validity, AVE[e]
		Unstd[a]	SE	Unstd/SE	*P* value	STD[b]	SMC[c]		
IU3	5.348	0.779	0.069	11.361	<.001	0.612	0.375		

[a]Unstd: unstandardized factor loadings.

[b]STD: standardized factor loadings.

[c]SMC: square multiple correlations.

[d]CR: composite reliability.

[e]AVE: average variance extracted.

[f]Not applicable.

Table 7. Discriminant validity.

Constructs	AVE[a]	PU	PR	IU	RB	EHC	SN	PEOU	PBC	TR
Perceived usefulness (PU)	0.431	*0.657*[b]	—[c]	—	—	—	—	—	—	—
Perceived risks (PR)	0.462	–0.266	*0.680*	—	—	—	—	—	—	—
Intention to use (IU)	0.506	0.458	–0.364	*0.711*	—	—	—	—	—	—
Resistance bias (RB)	0.524	–0.318	0.424	–0.374	*0.724*	—	—	—	—	—
Eye health consciousness (EHC)	0.625	0.309	–0.179	0.277	–0.272	*0.791*	—	—	—	—
Subjective norms (SN)	0.512	0.432	–0.289	0.471	–0.236	0.179	*0.716*	—	—	—
Perceived ease of use (PEOU)	0.516	0.430	–0.223	0.324	–0.244	0.171	0.296	*0.718*	—	—
Perceived behavioral control (PBC)	0.408	0.383	–0.360	0.374	–0.116	0.181	0.453	0.380	*0.639*	—
Trust (TR)	0.429	0.343	–0.332	0.422	–0.152	0.126	0.458	0.247	0.411	*0.655*

[a]AVE: average variance extracted.

[b]The items on the diagonal in italics represent the square root of the AVE; off-diagonal elements are the correlation estimates.

[c]Not applicable.

In Table 7, the square roots of the AVE values (the italic numbers on the diagonal) are higher than the numbers in the off-diagonal direction (correlations between a particular construct in the same column and other constructs in different rows) in the corresponding columns, indicating that the discriminant validity of all constructs meets the criteria of Fornell and Larcker [76].

Structural Model Analysis

Table 8 presents the model fit indicators with their respective criteria: (1) the standardized root mean square residual is 0.057, smaller than 0.08, (2) the comparative fit index is 0.915, greater than 0.90, and (3) the root mean squared error of approximation is 0.049, also smaller than 0.08. The model fit indicators shown in Table 8 satisfy most of the criteria and the combination rule [77], indicating that the hypothesized model has a good fit to the data.

Figure 2 shows the graphic description, and Table 9 shows the numerical results of the path coefficients. IU is significantly affected by SN (beta=.408; *P*<.001), PU (beta=.336; *P*=.03), and RB (beta=–.237; *P*=.02). PEOU (beta=.050; *P*=.59), EHC (beta=.077; *P*=.25), PBC (beta=–.066; *P*=.52) and PR (beta=–.133; *P*=.01) do not significantly affect IU. PEOU is significantly affected by PBC (beta=.506; *P*<.001) and SN (beta=.354; *P*=.002). PU is significantly affected by EHC (beta=.159; *P*<.001) and PEOU (beta=.279; *P*<.001).

R^2 was calculated to access the validity of the research model. As Table 9 and Figure 2 show, 51.5% of IU can be explained by PU, SN, PEOU, RB, and PR constructs; 48.8% of PU can be explained by the EHC and PEOU constructs; and 39.6% of PEOU can be explained by the SN and PBC constructs.

Table 8. Model fit of the research model.

Model fit	Criteria	Model fit of research model
χ^{2a}	The smaller the better	755.629
df	The larger the better	356.00
Normed chi-square (χ^2/df)	$1<\chi^2/df<3$	2.123
RMSEA[b]	<0.08	0.049
SRMR[c]	<0.08	0.057
CFI[d]	>0.9	0.915
GFI[e]	>0.9	0.896
AGFI[f]	>0.8	0.873

[a]χ^2: chi-square.
[b]RMSEA: root mean squared error of approximation.
[c]SRMR: standardized root mean square residual.
[d]CFI: comparative fit index.
[e]GFI: goodness-of-fit index.
[f]AGFI: adjusted goodness-of-fit index.

Figure 2. Estimates of regression analysis. Note: Solid line indicates a significant path and dotted line indicates a nonsignificant path.

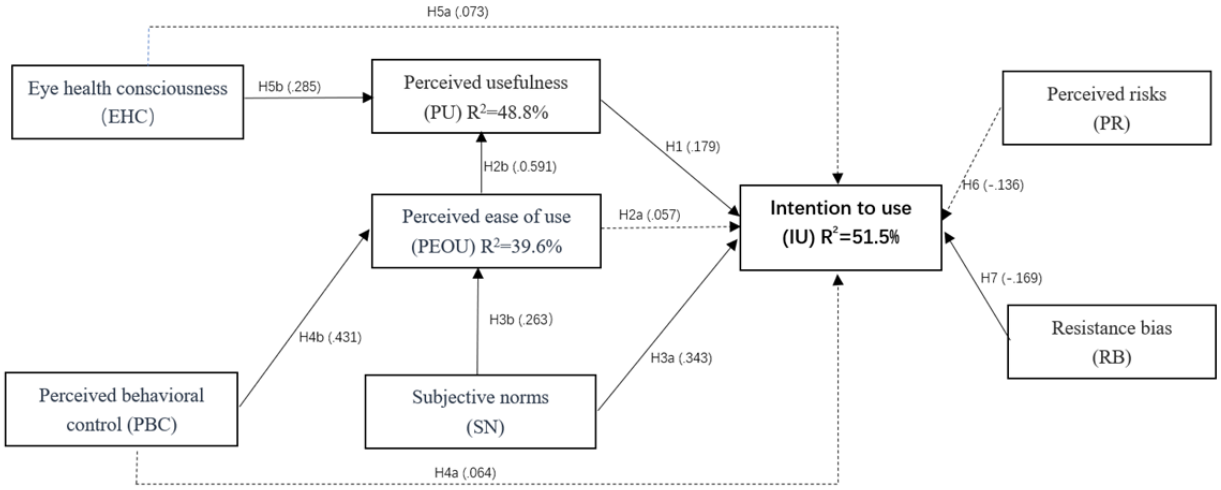


Table 9. Regression coefficient.

Dependent variables and hypothesis (H)	Unstd[a]	SE	*T* value	*P* value	Std[b]	Supported	R^2
IU[c]							0.515
IU←PU[d] (H1)	0.336	0.151	2.219	.03	0.179	✓	
IU←PEOU[e] (H2a)	0.05	0.093	0.544	.59	0.057	X	
IU←SN[f] (H3a)	0.408	0.098	4.146	<.001	0.343	✓	
IU←PR[g] (H6)	–0.124	0.066	–1.875	.06	–0.136	X	
IU←RB[h] (H7)	–0.237	0.102	–2.328	.02	–0.169	✓	
IU←EHC (H5a)	0.077	0.066	1.156	.25	0.073	X	
IU←PBC (H4a)	0.066	0.104	0.64	.52	0.064	X	
PU							0.488
PU←EHC[i] (H5b)	0.159	0.031	5.14	<.001	0.285	✓	
PU←PEOU (H2b)	0.279	0.034	8.128	<.001	0.591	✓	
PEOU							0.396
PEOU←SN (H3b)	0.354	0.116	3.051	.002	0.263	✓	
PEOU←PBC[j] (H4b)	0.506	0.11	4.59	<.001	0.431	✓	

[a]Unstd: unstandardized factor loadings.

[b]Std: standardized factor loadings.

[c]IU: intention to use.

[d]PU: perceived usefulness.

[e]PEOU: perceived ease of use.

[f]SN: subjective norms.

[g]PR: perceived risks

[h]RB: resistance bias.

[i]EHC: eye health consciousness.

[j]PBC: perceived behavioral control.

Analysis of Mediation Effects

Bias-corrected bootstrapping mediation analysis (5000 iterations) was used to examine the indirect effects (Table 10).

PU fully mediates the effect of EHC on IU (95% CI 0.04 to 0.1361). PEOU and PU fully mediate the effect of PBC on IU (95% CI 0.01 to 0.2322), whereas PEOU partially mediates the effect of PBC on PU (95% CI 0.057 to 0.2697). PEOU and PU do not mediate the effect of SN on IU (95% CI –0.0011 to 0.2000), whereas PEOU partially mediates the effect of SN on PU (95% CI 0.0004 to 0.2517). PU also fully mediates the effect of PEOU on IU (95% CI 0.005 to 0.2398).

Analysis of Moderation Effect

In Figure 3 and Table 11, the trust moderates the effect of PU on IU (beta=–.0.095; *P*=.049), where the effect of PU on UI is stronger for the users with low trust compared with those with high trust.

Table 10. Analysis of indirect effects.

Paths relationship	Direct effect (95% CI)			Indirect effect (95% CI)			Results
	Effect	LLCI[a]	ULCI[b]	Effect	LLCI	ULCI	
EHC[c]→PU[d]→IU[e]	0.0765	–0.0636	0.2443	0.053	0.004	0.1361	Fully
PBC[f]→PEOU[g]→PU→IU	0.0663	–0.202	0.3133	0.073	0.001	0.2322	Fully
PBC→PEOU→IU	0.0663	0.202	0.3133	0.073	0.001	0.2322	Fully
PEOU→PU→IU	0.0504	–0.1868	0.3059	0.094	0.005	0.2398	Fully
PBC→PEOU→PU	0	0	0	0.141	0.057	0.2697	Partial
SN[h]→PEOU→PU	0	0	0	0.099	0.0004	0.2517	Partial
SN→PEOU→PU→IU	0.4083	0.1768	0.6509	0.051	–0.0011	0.2	No
SN→PEOU→IU	0.4083	0.1768	0.6509	0.051	–0.0011	0.2	No

[a]LLCI: lower limit confidence interval.
[b]ULCI: upper limit confidence interval.
[c]EHC: eye health consciousness.
[d]PU: perceived usefulness.
[e]IU: intention to use.
[f]PBC: perceived behavioral control.
[g]PEOU: perceived ease of use.
[h]SN: subjective norms.

Figure 3. Trust moderates the effect of PU on IU. a *P*<.01; b *P*<.05.

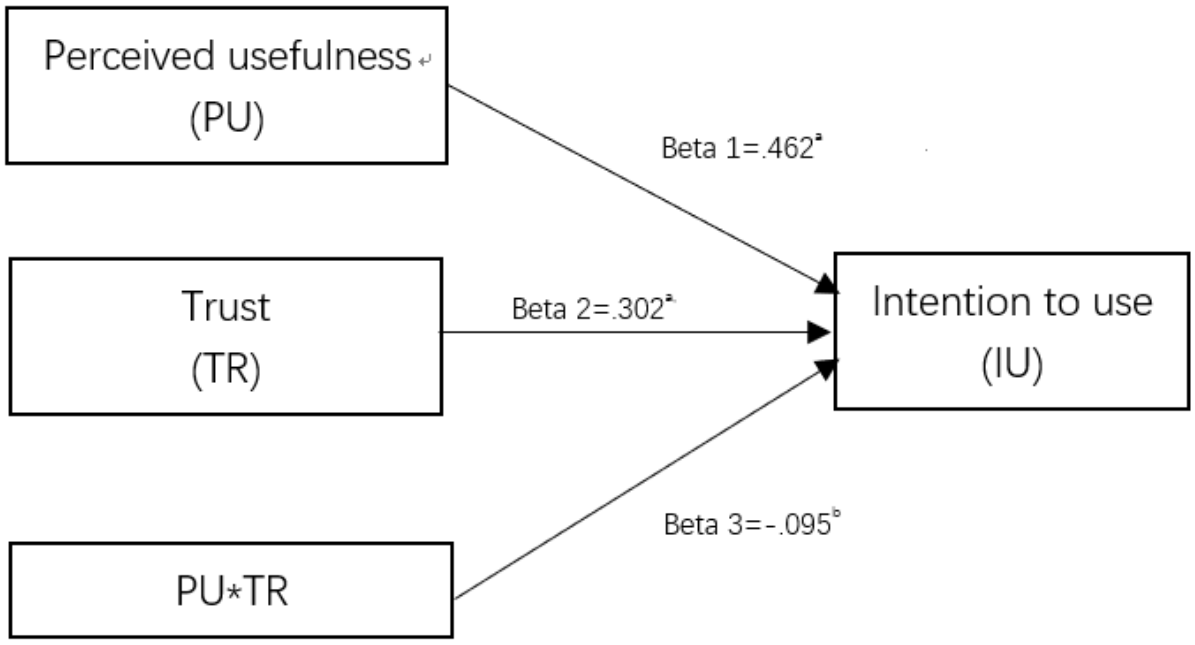


Table 11. Moderation analysis.

Dependent variable, independent variable	Unstd[a]	Std[b]	SE	*P* value	Bootstrap 1000 times, bias-corrected 95% CI
Intention to use					
Perceived usefulness	0.934	0.462	0.237	<.001	0.4691 to 1.3997
Trust	0.857	0.302	0.287	.003	0.2937 to 1.4209
Perceived usefulness×Trust	–0.095	–0.095	0.048	.049	–0.1897 to –0.0001

[a]Unstd: unstandardized factor loading.

[b]Std: standardized factor loadings

Discussion

Principal Findings

This study investigated the relationships between factors that affect the adoption of ophthalmic AI devices for eye health management. The research model was developed using relevant theories of technology acceptance, including TAM, TPB, HBM, DFT, SQB, and Trust to fit AI applications in particular health care scenarios in China.

There are 4 principal findings: (1) SN plays a more important role than PU through both direct and indirect paths; (2) RB of new technology reduces public IU of ophthalmic AI, whereas PR does not have an effect on public IU; (3) EHC and PBC have an indirect positive effect on the IU of AI through the mediators PU and PEOU; and (4) trust moderates the effect of PU on IU. The results are discussed in detail below.

Subjective Norms Play a Much More Important Role in Artificial Intelligence Adoption Than Perceived Usefulness

As many studies have discussed, PU, PEOU, SN, and PBC significantly influence IU [31,32,53]. However, the function of SN differs among cultures. Some studies have found no significant effects [29,78], whereas others have reported the opposite result [59]. In our study, SN was the most important predictor of IU, whose direct effect on IU was much stronger than that of PU. It also had a significant positive effect on PU through PEOU. These results indicate that, in China, when individuals encounter new technologies such as ophthalmic AI devices, public perceptions about usefulness, ease of use, and IU are likely to be influenced by their significant others (the items of the SN construct) such as close friends and relatives, colleagues and peers, and superiors or leaders in their work teams. This phenomenon could be linked to a crowd mentality (following the group's actions), collectivist culture (prioritizing a group over the individual), authoritarianism (follow the rule of team leaders), and Confucianism (conforming to prescribed relationship roles and avoiding transgression) in China.

Furthermore, an interesting finding was that PEOU did not have a significant direct effect on IU, so H2 was not supported. This finding means that the public's IU of ophthalmic AI was not influenced by perceptions of how easy these technologies would be to use. However, the average score of the construct PEOU was high, with a value of 5.739 out of 7. One possible explanation is that because ophthalmic AI devices are newly developed products, the public might perceive them as *intelligent* and believe that they should be convenient and easy to use. Although the direct effect of PEOU on IU was not significant, PEOU did have a strong effect on PU, confirming most TAM theories. Therefore, if someone whose opinion was important to participants suggested that they try the devices (SN), and the participants then realized the value or usefulness (PU) of the devices, participants' IU would be high.

Resistance Bias Reduces Public Intention to Use Ophthalmic Artificial Intelligence Whereas Perceived Risks Do Not Have an Effect on It

In most research on the Dual Factor Theory, PR negatively affected the public's IU. However, in our ophthalmic AI case, PR does not affect public's IU. This finding is in line with the Chinese context where people do not perceive risk of blindness as an acute threat and owing to the fact that the general population of China does not strongly prioritize privacy [79]. The low mean score of the 6 items of the PR construct (3.802 out of 7) reflects the public's lack of awareness of health risks and protection of health information and privacy. These results were also confirmed by our qualitative study that people are accustomed to providing key personal information when registering on an app or receiving nuisance calls.

We also integrated a new construct, RB, and verified its reliability and validity in our model, improving our understanding of negative factors involved in health care technology acceptance. Our results confirmed the SQB theory. People might reject ophthalmic AI devices owing to unfamiliarity, regret avoidance, or past experiences with new technology products. This resistance reflects many people's natural preference to continue with traditional approaches to health management. This finding matches observations about Chinese patients' acceptance of mobile phone health technology and mobile health services for chronic disease management that these inhibitors had a negative effect on behavioral intention [31,57].

Eye Health Consciousness and Perceived Behavioral Control Have an Indirect Positive Effect Via the Mediators Perceived Usefulness and Perceived Ease of Use

Previous studies of health behavior based on the theories of HBM have found that PHT (similar to EHC) has both direct and indirect effects on IU [31,36]. In our study, EHC had a significant positive influence on PU and an indirect influence on IU via PU. However, EHC had no significant direct effect on IU, which contrasts with the findings of Dou about Chinese

patients' acceptance of mobile phone health technology [31] but is consistent with the work of Kim about consumers' health behavior IU of health information technology [36]. Our findings could indicate that although people are conscious about their eye health and perceive health threats even without eye screening, they will assess the usefulness or function of new AI devices before switching from traditional face-to-face eye examination by ophthalmologists.

Many studies have found that perceived behavioral control has a significant direct and indirect influence on IU [53,57]. We found that it had no significant impact on IU, consistent with the meta-analysis of factors influencing mobile health service adoption [32], but in contrast with research on health professionals' adoption of health clouds [51] and physicians' acceptance of electronic medical record exchange [53]. Our findings could result from the different roles of general public and medical staff, as most health-related procedures in previous studies were conducted by medical staff, whose behavioral controllability of new developed devices was a more important concern during the manipulation process. However, we found that PBC had an indirect effect on IU through PEOU and PU. This indicates that unlike other health-related technologies studied, as emerging products, ophthalmic AI devices need to be convenient and useful to ensure the perception of behavioral controllability. The high average score of PBC items (5.005 out of 7) also shows that if the public perceived these devices as easy to use and useful, automanipulation of screening devices and self-management of eye screening could be achieved.

Moderation Effect of Trust

Previous studies have treated trust as a variable that affects IU directly or indirectly [44,52]. Few studies have discussed whether it could be a moderator. In China, in the context of unbalanced medical resource distribution and distrust between doctors and patients, this construct could play a more complicated role [66,68,80]. Our finding confirmed that this construct is a moderator, as trust had a significant moderation effect (beta=–.095; *P*=.049) on the path from PU to IU. The public's trust in the emerging technology and medical staff negatively moderated the influence of PU on IU. Participants with high trust in AI might have high expectations for AI in health care and thus might require greater PU before they would be willing to try the AI devices. Alternatively, participants with low trust in AI might have low expectations and require less PU before trying to use them. In light of the generally distrustful relationship between the public and medical staff in China, stakeholders such as doctors and AI suppliers should avoid making misleading or over-exaggerated claims in the promotion of AI health care products.

As the beta value was negative, the more the public trust AI devices, the lower the effect of PU on IU. In other words, if we improve people's beliefs and confidence about AI products, they will use these devices even if these devices are not as useful as they could be. In our study, the average score for the 3 trust items was 4.976 out of 7. Together with the low factor loading of PU on IU, in the Chinese context, the influence of PU was small. We interpreted this effect to be moderated by trust.

Comparison With Prior Work and Strengths of This Study

This study contributes to the AI health care literature in several ways:

1. This study was the first empirical study to examine the positive and negative factors that influence public acceptance of emerging AI devices in real clinical scenarios in China. As the model fit and R^2 values are high, our model can predict the Chinese public's IU of such AI devices.
2. We integrated 1 inhibitor of RB to modify the SQB theory to fit Chinese people's thinking style and language customs, and this showed both good convergence and discriminant validity.
3. We introduced trust as a moderator in the Chinese social context to reflect the health care context, and the results confirmed that trust had a moderation effect on the path from PU to IU.
4. SN has the greatest effect on the IU of AI devices. This finding differs from most studies of new health care technology acceptance and could reflect the culture, regulations, or rules in the Chinese social context.
5. PR does not significantly affect public's IU as participants were not aware of the protection of personal privacy and health information.

Implications for Practice

Researchers are only just beginning to assess how we might improve medicine using neural networks, and we will not know how well AI can predict key outcomes in health care settings without "robust validation in prospective, real-world clinical environments, with rigorous statistical methodology and analysis" [9]. As our data show, PU does not play as important a role as expected, and the following strategies could be a cost-effective way to improve public acceptance of AI devices and promote AI products in the era of narrow AI:

1. Enhance public trust in AI, avoiding misleading or exaggerated claims that might affect public perceptions of the function of AI in health care.
2. Expand the influence of SN through health communication campaigns in communities and workplaces, focusing on significant others in people's social circles such as superiors, public opinion leaders, and close friends.
3. Educate the public's knowledge and consciousness of the accuracy, effectiveness, safety, and privacy of AI devices and expedite legislation on AI to protect human rights.

Limitations and Future Research

Our nationwide study included people of all ages, from students to elders, which indicated good external validity. It was more cost-effective to recruit participants nationally from the internet rather than by traditional means, and this method was more suitable in the AI context because it provided real-time reports and feedback for target users. However, as the sample was collected through mobile devices or websites, the proportion of participants aged above 50 years was relatively low. This age distribution could reflect the fact that older people are less likely to use mobile devices owing to poor vision or motor abilities or RTC [58]. If automated or self-management procedures with

AI products require good mobile or digital skills, older people might not be appropriate target users. In future studies, we might introduce age as a moderator to evaluate its interference effect. When designing or promoting an AI device, we should consider its practical utility for older generations, as they are the main screening population in primary care projects. Moreover, medical staff such as hospital leaders, physicians, and nurses would be the main users of these devices, so their views on AI are very important. We will conduct further research on their intention to adopt and manipulate ophthalmic AI devices in real clinical scenarios.

Conclusions

Our study used the SEM method to explore the complex relationships between factors that influence public acceptance and IU of ophthalmic AI devices, as applied to real clinical scenarios in China. Positive factors such as SN played a more important and complex role than predicted, alongside people's EHC and PBC, whereas the inhibiting factor, RB, had a direct negative effect on adoption of AI devices. The new integrated inhibitor of RB fits Chinese people's thinking style and language customs and showed both good convergence and discriminant validity. PR does not significantly affect public's IU as they were not aware of the protection of personal privacy and health information. Furthermore, we found that trust had a moderation effect on the path from PU to IU. This integrated model, incorporating Chinese cultural and social contexts, demonstrated a good fit and explanatory power with high R^2 and could be used to explore other AI health care areas such as chronic disease screening and monitoring, especially for diabetes, hypertension, and cancer management.

Acknowledgments

This study was funded by the China Medical Board Open Competition program (grant number 18–299) and the PhD Start-up Fund of Natural Science Foundation of Guangdong Province of China (NO 2018A030310005). The authors would like to thank Prof Yucheng Liang, Prof Joseph D Tucker, and Willa Dong for their advice about the study. The authors especially thank Prof Yizhi Liu for sharing his knowledge on AI devices and their applications in clinical settings.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Answers for CHERRIES.
[PDF File (Adobe PDF File), 130 KB - jmir_v21i10e14316_app1.pdf]

References

1. Reddy S, Fox J, Purohit MP. Artificial intelligence-enabled healthcare delivery. J R Soc Med 2019 Jan;112(1):22-28. [doi: 10.1177/0141076818815510] [Medline: 30507284]
2. He J, Baxter SL, Xu J, Xu J, Zhou X, Zhang K. The practical implementation of artificial intelligence technologies in medicine. Nat Med 2019 Jan;25(1):30-36. [doi: 10.1038/s41591–018–0307–0] [Medline: 30617336]
3. Li X, Lu J, Hu S, Cheng K, de Maeseneer J, Meng Q, et al. The primary health-care system in China. Lancet 2017 Dec 9;390(10112):2584-2594. [doi: 10.1016/S0140-6736(17)33109-4] [Medline: 29231837]
4. Yang XH, Hu AL, Wang NL. From the prevention and treatment of blindness to the universal eye health. Ophthalmol CHN 2017;26(1):1-3. [doi: 10.13281/j.cnki.issn.1004-4469.2017.01.001]
5. Wang Y, Lv MZ, Zhu JF, He XG, He JN, Zou HD, et al. Fairness analysis of human resource allocation of primary eye care in shanghai city. Medicine and Society 2017;30(08):16-19. [doi: 10.13723/j.yxysh.2017.08.005]
6. Zhang G, Li Y, Teng X, Wu Q, Gong H, Ren F, et al. Prevalence and causes of low vision and blindness in Baotou: a cross-sectional study. Medicine (Baltimore) 2016 Sep;95(37):e4905 [FREE Full text] [doi: 10.1097/MD.0000000000004905] [Medline: 27631267]
7. Wang G, Bai Z, Shi J, Luo S, Chang HX, Sai X. Prevalence and risk factors for eye diseases, blindness, and low vision in Lhasa, Tibet. Int J Ophthalmol 2013;6(2):237-241 [FREE Full text] [doi: 10.3980/j.issn.2222-3959.2013.02.24] [Medline: 23638429]
8. Burlina PM, Joshi N, Pekala M, Pacheco KD, Freund DE, Bressler NM. Automated grading of age-related macular degeneration from color fundus images using deep convolutional neural networks. JAMA Ophthalmol 2017 Nov 1;135(11):1170-1176 [FREE Full text] [doi: 10.1001/jamaophthalmol.2017.3782] [Medline: 28973096]
9. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. Nat Med 2019 Jan;25(1):44-56. [doi: 10.1038/s41591–018–0300-7] [Medline: 30617339]
10. Kermany DS, Goldbaum M, Cai W, Valentim CC, Liang H, Baxter SL, et al. Identifying medical diagnoses and treatable diseases by image-based deep learning. Cell 2018 Feb 22;172(5):1122-31.e9 [FREE Full text] [doi: 10.1016/j.cell.2018.02.010] [Medline: 29474911]

11. Li Y, Huang W, Qiqige A, Zhang H, Jin L, Ti P, et al. Prevalence and causes of blindness, visual impairment among different ethnical minority groups in Xinjiang Uygur autonomous region, China. BMC Ophthalmol 2018 Feb 13;18(1):41 [FREE Full text] [doi: 10.1186/s12886–018–0705-6] [Medline: 29433477]
12. Keel S, Lee PY, Scheetz J, Li Z, Kotowicz MA, MacIsaac RJ, et al. Feasibility and patient acceptability of a novel artificial intelligence-based screening model for diabetic retinopathy at endocrinology outpatient services: a pilot study. Sci Rep 2018 Mar 12;8(1):4330 [FREE Full text] [doi: 10.1038/s41598–018-22612-2] [Medline: 29531299]
13. Long E, Lin H, Liu Z, Wu X, Wang L, Jiang J, et al. An artificial intelligence platform for the multihospital collaborative management of congenital cataracts. Nat Biomed Eng 2017 Jan 30;1(2):0024. [doi: 10.1038/s41551–016–0024]
14. China Big Data Industry Watch. 2018. Medical Artificial Intelligence Technology and Application White Paper (2018) URL: http://www.cbdio.com/BigData/2018–05/08/content_5711506.htm [accessed 2019–09–17]
15. Holden RJ, Karsh B. The technology acceptance model: its past and its future in health care. J Biomed Inform 2010 Feb;43(1):159-172 [FREE Full text] [doi: 10.1016/j.jbi.2009.07.002] [Medline: 19615467]
16. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quart 1989;13(3):319. [doi: 10.2307/249008]
17. Davis FD, Bagozzi RP, Warshaw PR. User acceptance of computer technology: a comparison of two theoretical models. Manag Sci 1989 Aug;35(8):982-1003. [doi: 10.1287/mnsc.35.8.982]
18. Schifter DE, Ajzen I. Intention, perceived control, and weight loss: an application of the theory of planned behavior. J Pers Soc Psychol 1985 Sep;49(3):843-851. [doi: 10.1037//0022-3514.49.3.843] [Medline: 4045706]
19. Ajzen I. The theory of planned behaviour is alive and well, and not ready to retire: a commentary on Sniehotta, Presseau, and Araújo-Soares. Health Psychol Rev 2015;9(2):131-137. [doi: 10.1080/17437199.2014.883474] [Medline: 26209198]
20. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. MIS Quart 2003;27(3):425. [doi: 10.2307/30036540]
21. Hu PJ, Chau PY, Sheng OR, Tam KY. Examining the technology acceptance model using physician acceptance of telemedicine technology. J Manage Inform Syst 1999;16(2):91-112. [doi: 10.1080/07421222.1999.11518247]
22. Chau PY, Hu PJ. Examining a model of information technology acceptance by individual professionals: an exploratory study. J Manage Inform Syst 2002;18(4):191-229. [doi: 10.1080/07421222.2002.11045699]
23. Rho MJ, Choi IY, Lee J. Predictive factors of telemedicine service acceptance and behavioral intention of physicians. Int J Med Inform 2014 Aug;83(8):559-571. [doi: 10.1016/j.ijmedinf.2014.05.005] [Medline: 24961820]
24. Sambasivan M, Esmaeilzadeh P, Kumar N, Nezakati H. Intention to adopt clinical decision support systems in a developing country: effect of physician's perceived professional autonomy, involvement and belief: a cross-sectional study. BMC Med Inform Decis Mak 2012 Dec 5;12:142 [FREE Full text] [doi: 10.1186/1472-6947–12–142] [Medline: 23216866]
25. Esmaeilzadeh P, Sambasivan M, Kumar N. The Challenges and Issues Regarding E-Health and Health Information Technology Trends in the Healthcare Sector. In: Proceedings of the International Conference on E-business Technology and Strategy. 2010 Presented at: CETS'10; September 29-30, 2010; Ottawa, Canada p. 23-37. [doi: 10.1007/978-3-642–16397-5_2]
26. Liu L, Ma Q. Perceived system performance: A test of an extended technology acceptance model. Data Base Adv Inf Sy 2006 Sep 19;37(2-3):51-59. [doi: 10.1145/1161345.1161354]
27. Lin C, Lin I, Roan J. Barriers to physicians' adoption of healthcare information technology: an empirical study on multiple hospitals. J Med Syst 2012 Jun;36(3):1965-1977. [doi: 10.1007/s10916–011-9656-7] [Medline: 21336605]
28. Egea JM, González MV. Explaining physicians' acceptance of EHCR systems: An extension of TAM with trust and risk factors. Comput Hum Behav 2011;27(1):319-332. [doi: 10.1016/j.chb.2010.08.010]
29. Tavares J, Oliveira T. New integrated model approach to understand the factors that drive electronic health record portal adoption: cross-sectional national survey. J Med Internet Res 2018 Nov 19;20(11):e11032 [FREE Full text] [doi: 10.2196/11032] [Medline: 30455169]
30. Wu JH, Wang SC, Lin LM. Mobile computing acceptance factors in the healthcare industry: a structural equation model. Int J Med Inform 2007 Jan;76(1):66-77. [doi: 10.1016/j.ijmedinf.2006.06.006] [Medline: 16901749]
31. Dou K, Yu P, Deng N, Liu F, Guan Y, Li Z, et al. Patients' acceptance of smartphone health technology for chronic disease management: a theoretical model and empirical test. JMIR Mhealth Uhealth 2017 Dec 6;5(12):e177 [FREE Full text] [doi: 10.2196/mhealth.7886] [Medline: 29212629]
32. Zhao Y, Ni Q, Zhou R. What factors influence the mobile health service adoption? A meta-analysis and the moderating role of age. Int J Inform Manage 2018;43:342-350. [doi: 10.1016/j.ijinfomgt.2017.08.006]
33. Liang H, Xue Y, Byrd TA. PDA usage in healthcare professionals: testing an extended technology acceptance model. Int J Mob Commun 2003;1(4):372-382. [doi: 10.1504/IJMC.2003.003992]
34. Yi MY, Jackson JD, Park JS, Probst JC. Understanding information technology acceptance by individual professionals: toward an integrative view. Inform Manag 2006;43(3):350-363. [doi: 10.1016/j.im.2005.08.006]
35. Pai F, Huang K. Applying the technology acceptance model to the introduction of healthcare information systems. Technol Forecast Soc Change 2011;78(4):650-660. [doi: 10.1016/j.techfore.2010.11.007]
36. Kim J, Park HA. Development of a health information technology acceptance model using consumers' health behavior intention. J Med Internet Res 2012 Oct 1;14(5):e133 [FREE Full text] [doi: 10.2196/jmir.2143] [Medline: 23026508]

37. Lai PC. The literature review of technology adoption models and theories for the novelty technology. J Inform Sys Technol Manag 2017;14(1):21-38. [doi: 10.4301/s1807–17752017000100002]
38. Ajzen I. From intentions to action: A theory of planned behavior. In: Kuhl J, Beckmann J, editors. Action Control: From Cognition to Behavior. New York: Springer; 1985:11-39.
39. Ajzen I. The theory of planned behavior. Organ Behav Hum Decis Process 1991;50(2):179-211. [doi: 10.1016/0749-5978(91)90020-t]
40. McEachan RR, Conner M, Taylor NJ, Lawton RJ. Prospective prediction of health-related behaviours with the Theory of Planned Behaviour: a meta-analysis. Heal Psychol Rev 2011 Sep;5(2):97-144. [doi: 10.1080/17437199.2010.521684]
41. Fishbein M, Ajzen I. Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research. Philippines: Addison Wesley; 1975.
42. Han H, Hsu LT, Sheu C. Application of the Theory of Planned Behavior to green hotel choice: testing the effect of environmental friendly activities. Tour Manag 2010;31(3):325-334. [doi: 10.1016/j.tourman.2009.03.013]
43. Venkatesh V, Davis FD. A theoretical extension of the technology acceptance model: four longitudinal field studies. Manag Sci 2000;46(2):186-204. [doi: 10.1287/mnsc.46.2.186.11926]
44. Safa NS, Von Solms R. An information security knowledge sharing model in organizations. Comput Hum Behav 2016;57:442-451. [doi: 10.1016/j.chb.2015.12.037]
45. Gumussoy CA, Calisir F. Understanding factors affecting e-reverse auction use: an integrative approach. Comput Hum Behav 2009;25(4):975-988. [doi: 10.1016/j.chb.2009.04.006]
46. Chen S, Chen H, Chen M. Determinants of satisfaction and continuance intention towards self‐service technologies. Ind Manag Data Syst 2009 Oct 30;109(9):1248-1263. [doi: 10.1108/02635570911002306]
47. Rosenstock IM. Historical origins of the health belief model. Heal Edu Monogr 1974;2(4):328-335. [doi: 10.1177/109019817400200403]
48. Janz NK, Becker MH. The Health Belief Model: a decade later. Health Educ Q 1984;11(1):1-47. [doi: 10.1177/109019818401100101] [Medline: 6392204]
49. Ahadzadeh AS, Sharif S, Ong FS, Khong KW. Integrating health belief model and technology acceptance model: an investigation of health-related internet use. J Med Internet Res 2015;17(2):e45 [FREE Full text] [doi: 10.2196/jmir.3564] [Medline: 25700481]
50. Cenfetelli RT, Schwarz A. Identifying and testing the inhibitors of technology usage intentions. Inform Syst Res 2011;22(4):808-823. [doi: 10.1287/isre.1100.0295]
51. Hsieh P. Healthcare professionals' use of health clouds: integrating technology acceptance and status quo bias perspectives. Int J Med Inform 2015 Jul;84(7):512-523. [doi: 10.1016/j.ijmedinf.2015.03.004] [Medline: 25842155]
52. Andrews L, Gajanayake R, Sahama T. The Australian general public's perceptions of having a personally controlled electronic health record (PCEHR). Int J Med Inform 2014 Dec;83(12):889-900. [doi: 10.1016/j.ijmedinf.2014.08.002] [Medline: 25200198]
53. Hsieh P. Physicians' acceptance of electronic medical records exchange: an extension of the decomposed TPB model with institutional trust and perceived risk. Int J Med Inform 2015 Jan;84(1):1-14. [doi: 10.1016/j.ijmedinf.2014.08.008] [Medline: 25242228]
54. Samuelson W, Zeckhauser R. Status quo bias in decision making. J Risk Uncertainty 1988;1(1):7-59. [doi: 10.1007/BF00055564]
55. Polites GL, Karahanna E. Shackled to the status quo: the inhibiting effects of incumbent system habit, switching costs, and inertia on new system acceptance. MIS Quarterly 2012;36(1):21-42. [doi: 10.2307/41410404]
56. Bhattacherjee A, Hikmet N. Physicians' resistance toward healthcare information technology: a theoretical model and empirical test. Eur J Inform Syst 2017;16(6):725-737. [doi: 10.1057/palgrave.ejis.3000717]
57. Deng Z, Mo X, Liu S. Comparison of the middle-aged and older users' adoption of mobile health services in China. Int J Med Inform 2014 Mar;83(3):210-224. [doi: 10.1016/j.ijmedinf.2013.12.002] [Medline: 24388129]
58. Guo X, Sun Y, Wang N, Peng Z, Yan Z. The dark side of elderly acceptance of preventive mobile health services in China. Electron Markets 2012;23(1):49-61. [doi: 10.1007/s12525–012–0112-4]
59. Yoon C. The effects of national culture values on consumer acceptance of e-commerce: online shoppers in China. Inform Manag 2009;46(5):294-301. [doi: 10.1016/j.im.2009.06.001]
60. Shareef MA, Kumar V, Kumar U, Dwivedi YK. e-Government Adoption Model (GAM): Differing service maturity levels. Gov Inform Quart 2011;28(1):17-35. [doi: 10.1016/j.giq.2010.05.006]
61. Heerink M, Kröse B, Evers V, Wielinga B. Assessing acceptance of assistive social agent technology by older adults: The Almere model. Int J of Soc Robotics 2010;2(4):361-375. [doi: 10.1007/s12369–010–0068-5]
62. Lee JD, See KA. Trust in automation: designing for appropriate reliance. Hum Factors 2004;46(1):50-80. [doi: 10.1518/hfes.46.1.50_30392] [Medline: 15151155]
63. Hengstler M, Enkel E, Duelli S. Applied artificial intelligence and trust—the case of autonomous vehicles and medical assistance devices. Technol Forecast Soc Change 2016 Apr;105:105-120. [doi: 10.1016/j.techfore.2015.12.014]
64. Dwivedi YK, Shareef MA, Simintiras AC, Lal B, Weerakkody V. A generalised adoption model for services: A cross-country comparison of mobile health (m-Health). Gov Inform Quart 2016;33(1):174-187. [doi: 10.1016/j.giq.2015.06.003]

65. Posey C, Lowry PB, Roberts TL, Ellis TS. Proposing the online community self-disclosure model: the case of working professionals in France and the U.K. who use online communities. Eur J Inform Syst 2010;19(2):181-195. [doi: 10.1057/ejis.2010.15]
66. Wang W, Zhang H, Washburn DJ, Shi H, Chen Y, Lee S, et al. Factors influencing trust towards physicians among patients from 12 hospitals in China. Am J Health Behav 2018 Nov 1;42(6):19-30. [doi: 10.5993/AJHB.42.6.3] [Medline: 30157998]
67. -. Violence against doctors: why China? Why now? What next? Lancet 2014 Mar 22;383(9922):1013. [doi: 10.1016/S0140-6736(14)60501-8] [Medline: 24656183]
68. Tucker JD, Cheng Y, Wong B, Gong N, Nie J, Zhu W, Patient-Physician Trust Project Team. Patient-physician mistrust and violence against physicians in Guangdong Province, China: a qualitative study. BMJ Open 2015 Oct 6;5(10):e008221 [FREE Full text] [doi: 10.1136/bmjopen-2015–008221] [Medline: 26443652]
69. Cuadrado E, Tabernero C, García R, Luque B, Seibert J. The role of prosocialness and trust in the consumption of water as a limited resource. Front Psychol 2017;8:694 [FREE Full text] [doi: 10.3389/fpsyg.2017.00694] [Medline: 28533760]
70. Bentler P, Chou C. Practical Issues in Structural Modeling. Sociological Methods & Research 1987;16(1):78-117. [doi: 10.1177/0049124187016001004]
71. Wenjuanxing. Wengjuanxin. 2019. Sample service URL: https://www.wjx.cn/sample/service.aspx.20190831 [accessed 2019–09–18]
72. Aggelidis VP, Chatzoglou PD. Using a modified technology acceptance model in hospitals. Int J Med Inform 2009 Feb;78(2):115-126. [doi: 10.1016/j.ijmedinf.2008.06.006] [Medline: 18675583]
73. Wei YL, Cao XQ, Cai YL, Zhai YK, Liu RF. Influencing factors of patients' willingness to use remote monitoring services forchronic diseases based on technology acceptance model. Chi Gen Pract 2018;21(08):958-964. [doi: 10.3969/j.issn.1007-9572.2017.00.214]
74. Anderson JC, Gerbing DW. Structural equation modeling in practice: a review and recommended two-step approach. Psychol Bull 1988;103(3):411-423 [FREE Full text] [doi: 10.1037//0033-2909.103.3.411]
75. Hair J, Ringle C, Sarstedt M. PLS-SEM: Indeed a Silver Bullet. Journal of Marketing Theory and Practice 2011;19(2):139-152. [doi: 10.2753/MTP1069-6679190202]
76. Fornell C, Larcker DF. Evaluating structural equation models with unobservable variables and measurement error. J Marketing Res 1981;18(1):39-50. [doi: 10.1177/002224378101800104]
77. Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. Struct Equ Modeling 1999;6(1):1-55. [doi: 10.1080/10705519909540118]
78. Adenuga KI, Iahad NA, Miskon S. Towards reinforcing telemedicine adoption amongst clinicians in Nigeria. Int J Med Inform 2017 Aug;104:84-96. [doi: 10.1016/j.ijmedinf.2017.05.008] [Medline: 28599820]
79. Wang H. Protecting Privacy in China: A Research on China's Privacy Standards and the Possibility of Establishing the Right to Privacy and the Information Privacy Protection Legislation in Modern China. Berlin, Heidelberg: Springer-Verlag; 2011.
80. Tucker JD, Wong B, Nie J, Kleinman A. Rebuilding patient–physician trust in China. Lancet 2016;388(10046):755-756. [doi: 10.1016/S0140-6736(16)31362-9]

Abbreviations

AI: artificial intelligence
AVE: average variance extracted
CHC: community health center
CR: composite reliability
DFT: Dual Factor Theory
DR: diabetic retinopathy
EHC: eye health consciousness
HBM: health belief model
IS: incumbent system
IU: intention to use
PBC: perceived behavioral control
PEC: primary eye care
PEOU: perceived ease of use
PHT: perceived health threat
PR: perceived risks
PU: perceived usefulness
RB: resistance bias
RTC: resistance to change
SEM: Structural equation modeling
SN: subjective norms

SQB: status quo bias
TAM: Technology Acceptance Model
TPB: Theory of Planned Behavior
TR: trust
UTAUT: Unified Theory of Acceptance and Use of Technology

Edited by G Eysenbach; submitted 12.04.19; peer-reviewed by J Tavares, FS Ong; comments to author 07.05.19; revised version received 31.08.19; accepted 02.09.19; published 17.10.19.

Please cite as:
Ye T, Xue J, He M, Gu J, Lin H, Xu B, Cheng Y
Psychosocial Factors Affecting Artificial Intelligence Adoption in Health Care in China: Cross-Sectional Study
J Med Internet Res 2019;21(10):e14316
URL: http://www.jmir.org/2019/10/e14316/
doi:10.2196/14316
PMID:31625950

Original Paper

Use of a Web-Based Dietary Assessment Tool (RiksmatenFlex) in Swedish Adolescents: Comparison and Validation Study

Anna Karin Lindroos[1,2], PhD; Jessica Petrelius Sipinen[1], MSc; Cecilia Axelsson[1], MSc; Gisela Nyberg[3], PhD; Rikard Landberg[4], PhD; Per Leanderson[5], PhD; Marianne Arnemo[1], BSc; Eva Warensjö Lemming[1], PhD

[1]Swedish National Food Agency, Uppsala, Sweden

[2]Department of Internal Medicine and Clinical Nutrition, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden

[3]Department of Public Health Sciences, Karolinska Institutet, Stockholm, Sweden

[4]Department of Biology and Biological Engineering, Chalmers University of Technology, Gothenburg, Sweden

[5]Occupational and Environmental Medicine Center, Department of Clinical and Experimental Medicine, Linköping University, Linköping, Sweden

Corresponding Author:
Anna Karin Lindroos, PhD
Swedish National Food Agency
Box 622
Uppsala, 75126
Sweden
Phone: 46 18175500
Fax: 46 18105848
Email: annakarin.lindroos@slv.se

Abstract

Background: A Web-based dietary assessment tool—RiksmatenFlex—was developed for the national dietary survey of adolescents in Sweden.

Objective: This study aimed to describe the Web-based method RiksmatenFlex and to test the validity of the reported dietary intake by comparing dietary intake with 24-hour dietary recalls (recall interviews), estimated energy expenditure, and biomarkers.

Methods: Adolescents aged 11-12, 14-15, and 17-18 years were recruited through schools. In total, 78 students had complete dietary information and were included in the study. Diet was reported a few weeks apart with either RiksmatenFlexDiet (the day before and a random later day) or recall interviews (face-to-face, a random day later by phone) in a cross-over, randomized design. At a school visit, weight and height were measured and blood samples were drawn for biomarker analyses. Students wore an accelerometer for 7 days for physical activity measurements. Dietary intake captured by both dietary methods was compared, and energy intake captured by both methods was compared with the accelerometer-estimated energy expenditure (EEest). Intake of whole grain wheat and rye and fruit and vegetables by both methods was compared with alkylresorcinol and carotenoid concentrations in plasma, respectively.

Results: The mean of the reported energy intake was 8.92 (SD 2.77) MJ by RiksmatenFlexDiet and 8.04 (SD 2.67) MJ by the recall interviews (P=.01). Intake of fruit and vegetables was 224 (169) g and 227 (150) g, and whole grain wheat and rye intake was 12.4 (SD 13.2) g and 12.0 (SD 13.1) g, respectively; the intakes of fruit and vegetables as well as whole grain wheat and rye did not differ between methods. Intraclass correlation coefficients ranged from 0.57 for protein and carbohydrates to 0.23 for vegetables. Energy intake by RiksmatenFlexDiet was overreported by 8% (P=.03) but not by the recall interviews (P=.53) compared with EEest. The Spearman correlation coefficient between reported energy intake and EEest was 0.34 (P=.008) for RiksmatenFlexDiet and 0.16 (P=.21) for the recall interviews. Spearman correlation coefficient between whole grain wheat and rye and plasma total alkylresorcinol homologs was 0.36 (P=.002) for RiksmatenFlexDiet and 0.29 (P=.02) for the recall interviews. Spearman correlations between intake of fruit and vegetables and plasma carotenoids were weak for both dietary tools. The strongest correlations were observed between fruit and vegetable intake and lutein/zeaxanthin for RiksmatenFlexDiet (0.46; P<.001) and for recall interviews (0.28; P=.02).

Conclusions: RiksmatenFlexDiet provides information on energy, fruit, vegetables, and whole grain wheat and rye intake, which is comparable with intake obtained from recall interviews in Swedish adolescents. The results are promising for cost-effective dietary data collection in upcoming national dietary surveys and other studies in Sweden. Future research should focus on how, and if, new technological solutions could reduce dietary reporting biases.

(J Med Internet Res 2019;21(10):e12572) doi:10.2196/12572

KEYWORDS

dietary assessment; 24-hour hour dietary recalls; internet; validity; biomarkers; carotenoids; alkylresorcinols; adolescents

Introduction

There is a need for cost-effective dietary assessment methods that provide data of high quality to enable studies on diet and health. Open-ended dietary assessment methods such as food diaries and interviewer-administered 24-hour dietary recalls are resource demanding. Food Frequency Questionnaires (FFQ) are therefore commonly used in large epidemiological studies to minimize costs and workload. FFQs, however, are less accurate than open-ended dietary assessment methods [1,2]. Furthermore, they provide less detail on the type and amounts of foods eaten and no information on meal patterns. Food consumption data need to be harmonized for dietary exposure assessments on a European level. To achieve harmonization, the European Food Safety Authority (EFSA) recommends two 24-hour dietary recall interviews with at least one face-to-face interview for national dietary surveys of adults and children from the age of 10 years [3].

In recent years, technological developments have made it possible to assess diet using online tools. This offers potential for cost savings in data collection [4], but biases related to the self-reported dietary intake may still be present [5]. Several computer and Web-based 24-hour dietary recall and food diary methods have been developed. In a recent review, 21 different 24-hour dietary recall tools for use in children and adults were identified [6]. Most of these tools were compared with 24-hour dietary recall interviews or food diaries, with generally good agreement between the methods [6]. Only a few Web-based 24-hour dietary recalls or food diaries have been validated against recovery biomarkers [7-11]. Several Web-based 24-hour dietary recall and food diary tools have been specifically developed or adapted for use in children and adolescents. These tools have been evaluated by comparison with 24-hour dietary recall interviews [12-16], food records [17], direct observation [12,14,18,19], estimated total energy expenditure based on accelerometer measurements [20,21], or concentration biomarkers [19,22,23].

Dietary assessment methods designed to characterize usual intake are difficult to validate, as their actual validity cannot be estimated with absolute certainty [24]. Ideally, dietary assessment methods should be validated against recovery biomarkers to allow assessment of their measurement errors and calibration. However, only a few recovery biomarkers are available. The doubly labeled water method (measuring energy expenditure) for validating energy intake is expensive, and measurement of nitrogen in 24-hour urine collections for protein intake requires very motivated study participants. An alternative option to validate energy intake is estimation of the total energy expenditure from accelerometer data and equations [20,21,25,26]. Although this estimation of energy expenditure is not without errors, the physical activity level is objectively measured. Thus, it can be assumed that errors are uncorrelated to subjectively reported energy intake. Another option is to use concentration biomarkers to evaluate intake of specific food groups. Concentration biomarkers do not capture total intake, but can be used to evaluate the intake of specific food groups, although the strength of agreement is expected to be lower than that for recovery biomarkers [27]. A high intake of whole grains, fruits, and vegetables is associated with a decreased risk of chronic disease [28], and it is therefore important to be able to measure the intake of these food groups correctly. Alkylresorcinols are found in the outer parts of the wheat and rye kernels and can be used as biomarkers for whole grain wheat and rye intake [29-31]. Carotenoids reflect intake of fruit and vegetables satisfactorily and have been used in several studies to validate fruit and vegetable intake [32,33].

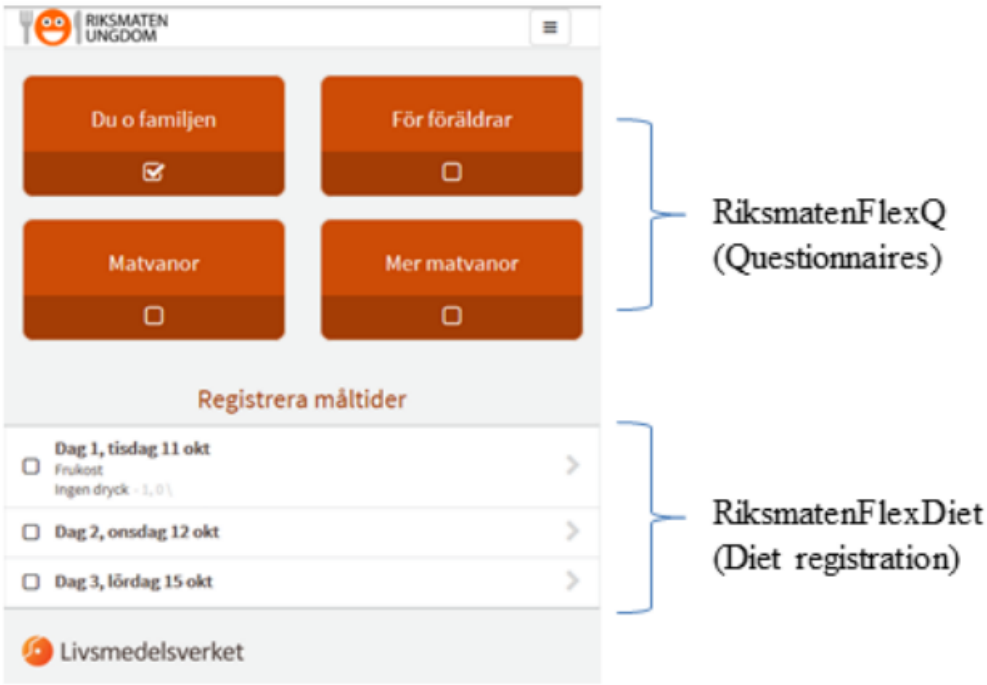
When planning the national dietary survey of adolescents in Sweden, there was an urgent need for a cost-effective, user-friendly dietary assessment method that could capture dietary intake satisfactorily and comply with EFSA's guidelines [3]. Therefore, a new Web-based dietary assessment tool—RiksmatenFlex—was developed for the survey of adolescents. The intention was also to provide a Web-based tool that could be further adapted to other age groups and study populations. As pointed out by Eldridge et al [4], new technology tools for assessing dietary intake require detailed publications describing the tool and testing the validity in order to meet general quality standards. Thus, the overall aim of this study is to describe the RiksmatenFlex tool and evaluate the validity of the dietary registration part of the tool in adolescents. The validity was evaluated by (1) comparing the reported energy and macronutrient intake with interview-administered 24-hour dietary recalls; (2) comparing the reported energy intake with total energy expenditure estimated from weight, height, and accelerometer data; (3) comparing the intake of whole grain wheat and rye, fruits, and vegetables with the plasma concentration of biomarkers alkylresorcinols and carotenoids, respectively; and (4) evaluating performance of the tool in (2) and (3) with the performance of interviewer-administered 24-hour dietary recalls.

Methods

The RiksmatenFlex Tool

The RiksmatenFlex is a self-administered, Web-based method for use on smartphones, tablets, and computers. It includes a diet registration part (RiksmatenFlexDiet) and a questionnaire part (RiksmatenFlexQ; Figure 1).

Figure 1. The start page of RiksmatenFlex.



RiksmatenFlex was developed in 2014-2015 by the National Food Agency in Sweden (NFA) for the national dietary survey of adolescents. To ensure a user-friendly and attractive system, an expert in interaction design was also involved in the development in addition to experts in nutrition and information technology. The development was based on experiences from a Web-based food diary [8,34,35] and an interview study with adolescents (data not published). The process also included focus groups throughout the development and test sessions with adolescents. The system was Windows-based, and in 2015, the following development environment was used: SQL server 2014, Windows Server 2012, MVC 5, .NET 4 and .NET 4.5, and VS2013-17. The system is compatible with IE 9-11, Chrome, Win 7, and Office 2010 or later versions.

The Web page is accessed by individual usernames and passwords. At the first log in, participants provide their email and phone number for future communication. At this time, the dates for the registration days are also generated. RiksmatenFlexQ is described in more detail elsewhere [36].

RiksmatenFlexDiet

Registration of Food and Drinks

The dietary registration is based on the 24-hour dietary recall method. However, the method is flexible and can be used as a food diary. It can also be easily adapted for new studies and age groups other than adolescents. The different steps of RiksmatenFlexDiet are displayed in Swedish in Figure 2. In brief, the steps include (1) selection of the required recording day; (2) selection of the time of the eating/drinking occasion; (3) selection of the type of meal (breakfast, lunch, dinner/supper, snack, other eating, or drink only) and place of the eating/drinking occasion (at home, at school, in a restaurant/bar/café, at an event such as cinema/theatre/sports, someone else's home, fast food restaurant, other place, on the way car/bus/train, or at work); (4) search for foods and drinks using a built-in search engine linked to a food list (see below); and (5) selection of portion size and, where relevant, specification of the details of the dish (for example, type of meat in a casserole).

Figure 2. Overview of the different dietary recording steps in RiksmatenFlexDiet.

To aid participants in finding the correct food, different spelling options and brand names were included in the search engine. Furthermore, pictures of foods commonly consumed in the five food categories—bread, ready-to-eat sandwiches, breakfast cereals, ice cream, and fat spreads—are linked to the search engine (Multimedia Appendix 1). The program uses machine learning to adapt the search list to user preferences.

RiksmatenFlexDiet is self-instructive and includes several logical reminders, for example, a reminder to help participants remember to register drinks. Further, before submitting the completed registration, participants are presented with a list of commonly forgotten foods and are asked to review all recorded foods and drinks. Participants are also asked to specify if the recording day was an ordinary day, not an ordinary day, or an ill day. Automatic reminders were sent to participants if a registration day was not completed. Time to complete one day's tasks was 15-30 minutes, depending on age. All recorded foods and drinks, together with portion sizes, are directly stored in a database and automatic calculation of energy, nutrient, and food group intakes is enabled through direct link to the food composition database. The data output is flexible, and it is possible to extract foods from composite dishes and raw agricultural commodities or obtain intake by day, meal type, meal place, and time of the intake. All foods are FoodEx2 coded.

The Food List

The food list compiled for RiksmatenFlexDiet is built on the Swedish national food composition database with the aim to represent Swedish adolescents' food consumption. Foods were selected based on the amounts and frequencies of consumed foods, drinks, and dishes reported in a previous survey of adults [37] and a study of children aged 11-12 years (data not published). Key foods were also identified in focus groups with adolescents. To facilitate the search in RiksmatenFlexDiet, the number of foods was limited to approximately 800. Thus, some of the foods are generic, and sales statistics from commercial companies were used to compile the composition of these foods. For many generic foods, it is also possible to specify details. For example, the type of meat (pork, beef, etc) and liquid base (tomato, cream, etc) can be specified for a meat casserole (Figure 2-5a). However, brand names are not included in the specifications. This validation study included 761 core foods, but with the additional details, there were approximately 2300 possible food item combinations available. The food list is presented in Multimedia Appendix 2.

Picture Portion Guide

The picture portion guide included in RiksmatenFlexDiet is an extended version of the portion guide developed for the national dietary survey Riksmaten 2010-2011 [37]. The guide includes household measures, pieces, portion pictures of glasses and cups, and photos of 39 different food categories, with four to

eight different reference sizes in each category Figure 2-5b. Six reference sizes were most common.

The Validation and Comparison Study

Study Population

The study population consists of schoolchildren taking part in a feasibility study preceding the national dietary survey Riksmaten Adolescents 2016-2017.

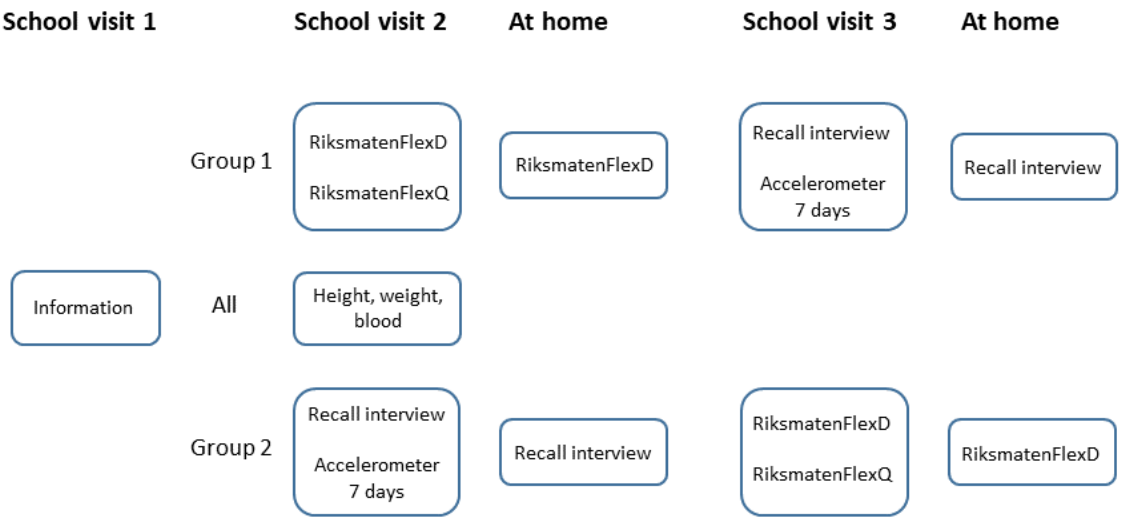
We aimed for a sample size of 75 adolescents, based on the assumption that more than 50 participants are sufficient when validating dietary intake against biomarkers [38]. Participants were recruited from elementary and high schools in grades 5, 8, and 11 across Sweden. Schools were selected from the Swedish school unit register with the aim to recruit one class from each school grade and Occupational and Environmental Medicine Centre (OEMC) region. In total, 86 schools were contacted, 15 accepted participation, and 18 classes (one class in each grade and region) were selected to participate. All students in the 18 selected classes were invited to participate in the study. Exclusion criterion was not being able to read and write Swedish. For participants to be included in the validation analyses, they needed to have completed 2 days of diet registration/recall with both methods.

When a school agreed to take part in the study, information letters including consent forms were sent to the eligible participants as well as the parents of the students in grades 5 and 8 of the selected class. All participants consented to take part, and for children younger than 16 years, their legal guardians provided written consent. Participants were given a gift voucher of 300 Sk (Swedish krona). The Regional Ethical Review Board in Uppsala, Sweden, approved the study (2015/190).

Study Design

The RiksmatenFlexDiet was compared with 24-hour dietary recall interviews (recall interviews), and diet by both methods was compared with estimated energy expenditure and biomarkers. As interviews about food intake might influence how well the diet is later reported in the self-reported RiksmatenFlexDiet and because blood was only drawn at school visit 2, the participants were randomized to either start with RiksmatenFlexDiet or recall interviews. The randomization was performed at the individual level for one class at a time using a simple randomization procedure. Figure 3 provides an outline of the study design. The study started with an information session at school with the selected classes (visit 1) when the study staff informed the students about the study, and the students had the possibility to ask questions. At the first examination day (visit 2), the signed consent forms were collected by the study staff and the participants started with either RiksmatenFlexDiet or a recall interview. At visit 3, the participants changed to the other method. Time between school visits 2 and 3 was 2-4 weeks. For both methods, diet was reported on a random day at home, 1-2 weeks after the school visits. Weight and height were measured and blood was drawn at school visit 2. Trained research assistants recruited the classes, informed the schools and teachers about the study, and carried out the fieldwork. Phlebotomists from six of the seven OEMCs in Sweden carried out the blood sampling.

Figure 3. Study design of the validation study of RiksmatenFlexDiet. RiksmatenFlexD: RiksmatenFlexDiet.



RiksmatenFlexDiet

RiksmatenFlexDiet was used both as a self-administered 24-hour dietary recall and as a food diary depending on day and age. The 24-hour dietary recall approach was used for the first day for all participants and for the second day for participants in grade 11. For the younger age groups (grade 5 and grade 8), the second day was recorded prospectively, as the interview study preceding the development of RiksmatenFlexDiet had indicated that they might have difficulties remembering yesterday's intake. Day 1 intake was recorded at school and day 2 intake was recorded at home and randomly assigned to occur 7-14 days

later. The first-day recording usually occurred during the school week, and around 80% of the second-day recordings were generated as a weekend day including Fridays. The date was unknown until the participants received an automated email with instructions to register food intake. Automatic email reminders were sent to participants if they did not complete their registrations. All records with an energy intake less than 800 kcal or over 3500 kcal were manually checked. Records with an intake below 800 kcal were judged as incomplete and excluded if no illness, less than three intake occasions (defined as at least one energy-contributing food item), and no meal after 3 PM was recorded (three records identified, one excluded). Records with an intake of over 3500 kcal were checked for obvious misunderstandings of amounts and foods (three records identified, none excluded).

The intake of energy, macronutrients, beta-carotene, and total whole grains was calculated using the NFA food composition database (the Riksmaten adolescent pilot study; the foods are displayed in Multimedia Appendix 2). Whole grain wheat and rye intake was calculated by excluding whole grains from sources other than wheat and rye in the total whole grain intake. Fruit and vegetable intake included all fruit (including berries) and vegetables (not potatoes and pulses) consumed, including both amounts directly recorded and amounts extracted from dishes. Juice was excluded from the intake due to the different nutrient density.

Recall Interviews

On day 1, a face-to-face interview was conducted in school and on day 2, the interview was performed over the phone. The second day was randomly assigned 7-14 days after the first interview, taking the weekday/weekend day into account as was done for RiksmatenFlexDiet. After the first interview, an appointment for the second phone interview was made with the participant. Trained staff with a nutritional background carried out all interviews, and the interviews followed a 24-hour dietary recall multiple pass protocol. The interview covered the period between waking up on the preceding day and waking up on the interview day. Household measures and a Web-based portion guide based on the pictures included in RiksmatenFlexDiet were used to describe the amounts of foods consumed. Participants were also asked if the previous day was a normal day. After the interview, the reported intake was coded using the nutrient calculation program Dietist Net Pro (Kost och Näringsdata AB, Stockholm, Sweden) and the NFA food composition database (recall interviews). Intake of whole grain wheat and rye, fruits, and vegetables was calculated as was done for RiksmatenFlexDiet.

Background Information

Educational levels of both parents were reported in RiksmatenFlexQ [36]. The information was combined and dichotomized to at least one parent with postsecondary education or no parent with postsecondary education.

Anthropometric Measurements and Blood Sampling

Height and weight were measured and nonfasting blood was collected from all participants at visit 2. Height was measured to the nearest 0.1 cm without shoes and weight, to the nearest 0.1 kg in light indoor clothing. Body mass index was calculated, and overweight/obesity was defined according to the recommendations by the International Obesity Task Force [39]. Height and weight were also used to estimate energy expenditure (see below).

Estimated Energy Expenditure

Physical activity was measured with an accelerometer (ActiGraph GT3X or wGT3X+, Tri-axis Accelerometer Monitor; Actigraph LLC, Pensacola, Florida). The participants were given the accelerometer after the first recall interview and instructed to wear it from waking up in the morning until bedtime for 7 consecutive days. The accelerometer was attached to an elastic band and placed on the right hip. Participants were instructed to remove the accelerometer when showering or doing water activities. Data were analyzed using ActiLife 6 Data Analysis Software (Actigraph LLC). Measurements of at least 3 complete days, including one weekend day (Saturday and Sunday), were considered as complete. Physical activity was assessed between 7 AM and 11 PM. The epoch length was set to 5 seconds. A complete day includes 500 minutes of registration. Based on mean counts per minute (cpm), physical activity energy expenditure was estimated using the modified prediction equations by Ekelund et al [40]:

$$\text{Physical activity energy expenditure (kJ/day)} = 4.182\times(66.847+(\text{cpm}\times 0.953)-(176.91\times\text{sex})$$

Total energy expenditure was calculated as physical activity energy expenditure+basal metabolic rate+diet-induced thermogenesis, where diet-induced thermogenesis was estimated as 10% of the total energy expenditure [41]. The basal metabolic rate was estimated from the Henry equations based on age, sex, height, and weight [42].

Biochemical Analyses

Alkylresorcinols and carotenoids were analyzed in plasma. Five participants did not provide enough blood for the analyses. In addition, one alkylresorcinol analysis failed.

Gas chromatography-mass spectrometry (GC-MS) was used to quantify the alkylresorcinols C17:0-C25:0 in 0.2 mL plasma samples, as described previously [43]. The intra- and interassay coefficients of variation were both 15% for total alkylresorcinols.

Plasma concentrations of carotenoids were determined with high-performance liquid chromatography, as previously described [34]. The intra- and interassay coefficients of variation varied between 5% and 7%.

Statistical Analysis

STATA statistical software (Version 14.2; STATA Corp, College Station, Texas) was used for the statistical analyses. Background information is presented as proportions and medians with the first and third quartiles. A mean of data from 2 days was used for all the analyses of reported dietary intake, and descriptive data are presented as mean and SD. Paired *t* test was used to test the difference in intake between the methods. Bland-Altman plots were used to assess precision and bias between reported intake by both methods and between reported

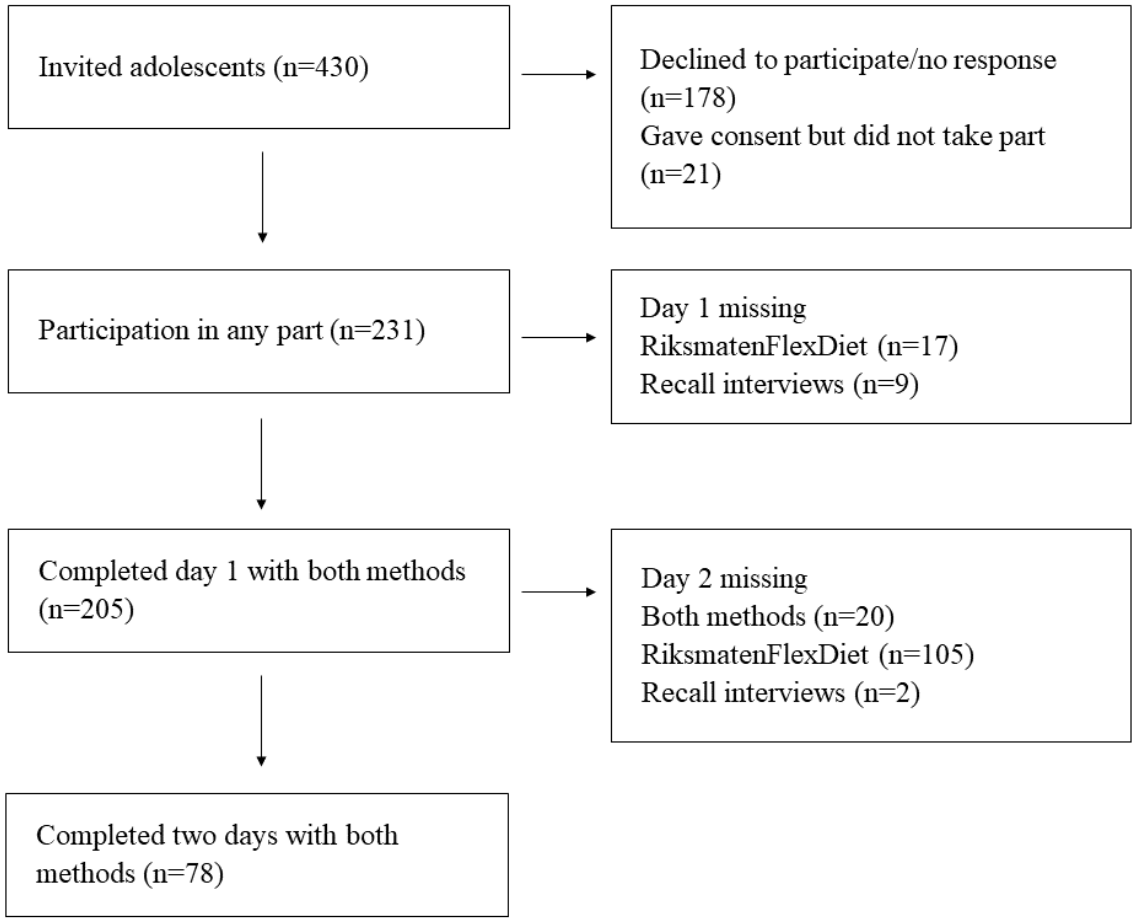
energy intake by the respective method and estimated energy expenditure. Agreement for reported energy and macronutrient intake by the two methods was calculated by tertiles, and a linear weighted kappa was used to evaluate the agreement between classifications. Intraclass correlations between the two dietary methods were calculated using a two-way mixed-effects model to evaluate the strength of agreement. Energy-adjusted (g/MJ) intake of fruit and vegetables and whole grain wheat and rye were correlated with the respective biomarkers by calculating the Spearman rank correlation coefficients.

Results

Study Population and Reported Days

The participant flow is illustrated in Figure 4. A total of 430 adolescents were invited to participate, of which 397 were randomized to either start with RiksmatenFlexDiet (n=194) or recall interviews (n=203). In total, 231 (54%) participated in at least one part of the study, 205 participants completed the first day by both methods (113 started with RiksmatenFlexDiet and 118 started with recall interviews), and 78 participants completed 2 days by both methods (36 started with RiksmatenFlexDiet and 42 started with recall interviews). Eighteen of the participants with complete diet information did not have complete information from the accelerometer measurements.

Figure 4. Participant flow of the validation study of RiksmatenFlexDiet.



Characteristics and plasma biomarkers of the 78 participants with complete diet information are presented in Table 1. Girls were more likely to participate than boys. The 60 participants with complete diet and accelerometer data were similar to the 78 participants with complete diet information. Lycopene was the most abundant carotenoid (36%), and the ratio of alkylresorcinol C17:0/C21:0 was 0.16.

The number of weekdays and weekend days (Friday-Sunday) was 93 (60%) and 63 (40%), respectively, for RiksmatenFlexDiet and 96 (62%) and 60 (38%), respectively, for recall interviews, which was close to the expected distribution for both methods.

Table 1. Characteristics, plasma concentrations, and counts per minute for participants with complete diet information (n=78).

Variables	Values
Sex, n (%)	
Girls	54 (69)
Boys	24 (31)
Academic school year (years), n (%)	
Grade 5 (11-12 years)	23 (29)
Grade 8 (14-16 years)	33 (44)
Grade 11 (17-18 years)	22 (28)
Parental education, n (%)	
Higher level[a]	39 (50)
Lower level	30 (38)
Missing information	9 (12)
Overweight, obese[b], n (%)	15 (19)
Plasma carotenoid concentration (µmol/L, n=73), median (p25, p75)[c]	
Alpha-carotene,	0.19 (0.13, 0.38)
Beta-carotene	0.79 (0.52, 1.04)
Beta-cryptoxantin	0.14 (0.11, 0.19)
lutein+zeaxantin	0.36 (0.26, 0.47)
Lycopene	0.86 (0.71, 1.06)
Total carotenoids, median (p25, p75)	2.39 (1.97, 3.03)
Plasma alkylresorcinol concentration (nmol/L, n=72), median (p25, p75)	
C:17	3.2 (1.3, 6.1)
C:19	15.8 (9.0, 32.3)
C:21	24.8 (12.9, 48.5)
C:23	6.1 (2.4, 11.7)
C:25	3.7 (1.5, 8.2)
Total alkylresorcinols	51.1 (29.9, 101.6)
C17:0/C21:0	0.16 (0.08, 0.25)
Physical activity (counts/min, n=60), median (p25, p75)	376 (289, 492)

[a]At least one parent had postsecondary education.

[b]Based on Cole and Lobstein [39].

[c]25th and 75th percentiles.

Reported Intake by RiksmatenFlexDiet and Recall Interviews

Reported intakes of energy, macronutrients, beta-carotene, whole grain wheat and rye, vegetable and fruit by RiksmatenFlexDiet and recall interviews are presented in Table 2. Reported energy and macronutrient intake was higher with RiksmatenFlexDiet than with recall interviews. However, intake of beta-carotene, whole grain wheat and rye, vegetables, and fruits did not differ between the methods.

Table 2. Reported dietary intake by RiksmatenFlex and recall interviews (n=78).

Dietary intake/day[a]	RiksmatenFlexDiet, mean (SD)	Recall interviews, mean (SD)	*P* value[b]	Intraclass correlation (95% CI)
Energy, MJ	8.92 (2.77)	8.04 (2.67)	.01	0.53 (0.35-0.67)
Protein, g	85 (33)	74 (24)	<.001	0.57 (0.39-0.70)
Fat, g	86 (30)	76 (33)	.01	0.27 (0.05-0.47)
Carbohydrates, g	243 (85)	225 (78)	.04	0.57 (0.40-0.70)
Dietary fiber, g	20 (10)	17 (6.0)	.01	0.45 (0.26-0.61)
Whole grain wheat and rye, g	12.4 (13.2)	12.0 (13.1)	.82	0.29 (0.07-0.48)
Beta-carotene, mg	1975 (2415)	1876 (2359)	.71	0.50 (0.32-0.65)
Vegetables, g	137 (106)	139 (95)	.90	0.23 (0.01-0.43)
Fruits, g	87 (112)	88 (98)	.89	0.56 (0.39-0.70)
Fruits and vegetables, g	224 (169)	227 (150)	.88	0.49 (0.30-0.64)

[a]Mean of two days

[b]Paired *t* test.

To investigate the agreement between the two methods, Bland-Altman plots were used for energy and macronutrient intake (Multimedia Appendix 3). For all comparisons, a few outliers were noted, but the majority fell within the limits of agreement. The intraclass correlations between the two methods were moderate and ranged from 0.57 for protein to 0.23 for vegetables. The strength of agreement for reported energy intake and macronutrient intake was fair to moderate. For energy, 87% were classified into the same or the adjacent tertile, and the weighted kappa was 0.42 (95% CI 0.25-0.58). Corresponding figures for protein, fat, and carbohydrates were 91% and 0.36 (95% CI 0.19-0.53), 86% and 0.22 (95% CI 0.04-0.39), 96% and 0.51 (95% CI 0.36-0.66), respectively.

Reported Energy Intake and Estimated Energy Expenditure

Mean accelerometer-estimated energy expenditure was 8.13 MJ (95% CI 7.80-8.46) compared with the mean reported energy intake of 8.80 (95% CI 8.19-9.42) and 8.05 (7.36-8.74) by RiksmatenFlexDiet and Recall interviews, respectively. Estimated energy expenditure was overestimated by RiksmatenFlexDiet (*P*=.02) but not by the recall interviews (*P*=.82). The agreement between reported energy intake and estimated energy expenditure for the two methods was also examined in Bland-Altman plots (Figures 5 and 6). The limits of agreement were wider for the comparison with recall interviews than for RiksmatenFlexDiet. Both plots show a bias toward overestimation at higher means. The Spearman correlation between estimated energy expenditure and reported energy intake was 0.34 (*P*=.008) for RiksmatenFlexDiet and 0.16 (*P*=.21) for recall interviews.

Reported Intake Versus Biomarkers

The Spearman correlations between the alkylresorcinol concentrations and reported intake of whole grain wheat and rye by both dietary methods were moderate but statistically significant for all comparisons (Table 3). The correlations were consistently higher for RiksmatenFlexDiet than for recall interviews.

The correlations between carotenoids and intake of vegetables and fruit were weak for both methods (Table 4). The strongest correlation was observed for lutein/zeaxanthin and intake of fruit and vegetables for RiksmatenFlexDiet (r=0.46; *P*<.001).

Figure 5. Bland-Altman plot of reported energy intake by RiksmatenFlexDiet and estimated energy expenditure (n=60). EEest: estimated total energy expenditure.

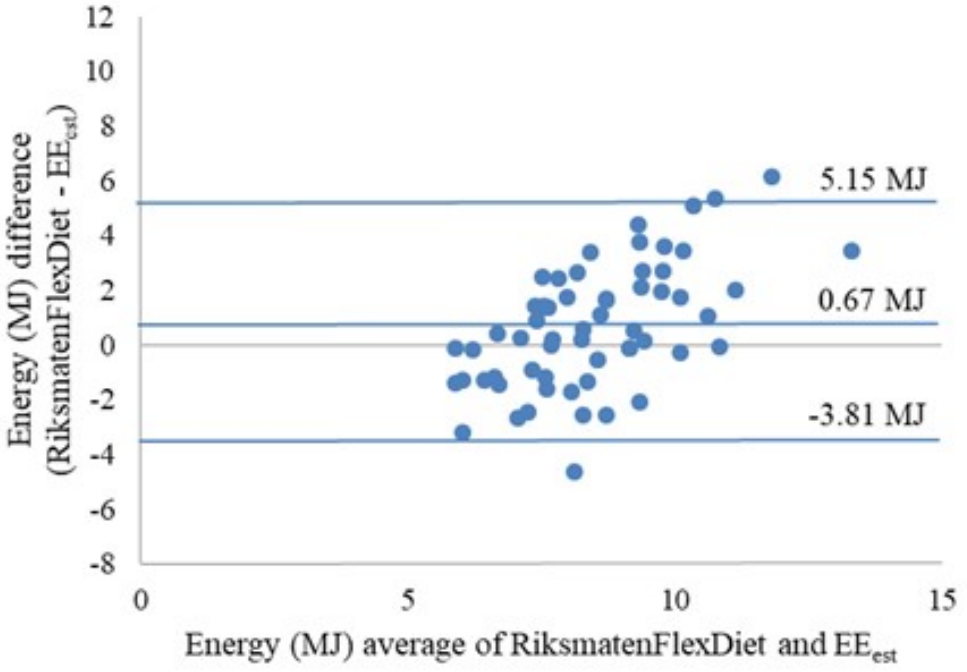


Figure 6. Bland-Altman plot of energy intake by recall interviews and estimated energy expenditure (n=60). EEest: estimated total energy expenditure.

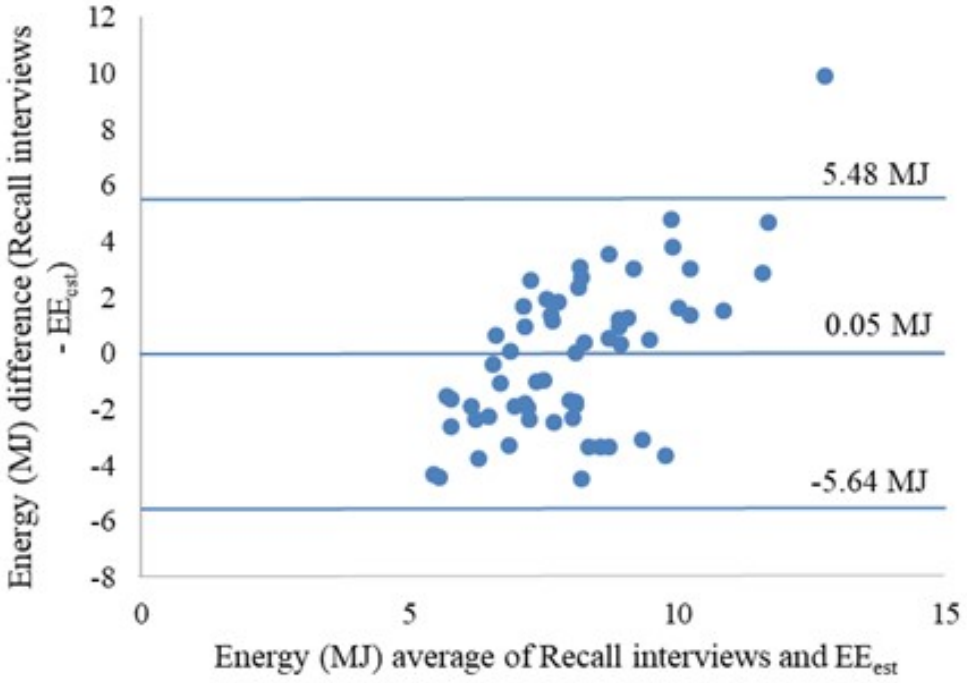


Table 3. Spearman rank correlations between energy-adjusted intake (g/MJ) of whole grain wheat and rye and plasma alkylresorcinol concentrations (n=72).

Alkylresorcinol	RiksmatenFlexDiet		Recall interviews	
	ρ[a]	*P* value	ρ	*P* value
C:17	0.43	<.001	0.32	.006
C:19	0.37	.001	0.29	.01
C:21	0.31	.008	0.26	.26
C:23	0.34	.004	0.27	.02
C:25	0.37	.001	0.29	.01
Total alkylresorcinols	0.36	.002	0.29	.01

[a]Spearman rank correlation coefficient.

Table 4. Spearman rank correlations between energy-adjusted intake (g/MJ) of fruit and vegetables and beta-carotene and plasma carotenoid concentrations (n=73).

Carotenoid	Vegetables		Fruits		Fruits and vegetables		Beta-carotene	
	ρ[a]	*P* value	ρ	*P* value	ρ	*P* value	ρ	*P* value
RiksmatenFlexDiet								
Alpha-carotene	–0.04	.72	0.22	.06	0.09	.47	0.28	.01
Beta-carotene	–0.01	.95	0.15	.20	0.07	.58	0.22	.06
Beta-cryptoxanthin	0.00	.98	0.17	.14	0.11	.36	0.12	.30
Lutein/zeaxanthin	0.37	.001	0.28	.02	0.46	<.001	0.16	.19
Lycopene	0.28	.02	0.05	.68	0.21	.08	0.25	.03
Total carotenoids	0.15	.21	0.15	.21	0.19	.11	0.32	.006
Recall interviews								
Alpha-carotene	–0.19	.11	0.21	.08	–0.01	.95	0.28	.02
Beta-carotene	–0.06	.59	0.08	.49	0.02	.90	0.26	.03
Beta-cryptoxanthin	–0.10	.39	0.06	.60	–0.06	.61	–0.03	.80
Lutein/zeaxanthin	0.29	.01	0.24	.04	0.28	.02	0.03	.81
Lycopene	0.17	.15	0.01	.91	0.11	.36	0.07	.56
Total carotenoids	–0.00	.97	0.12	.33	0.06	.62	0.21	.07

[a]Spearman rank correlation coefficient.

Discussion

Principal Findings and Comparison With Previous Work

The results of this study suggest that RiksmatenFlexDiet provides dietary intake estimates comparable to recall interviews in adolescents. RiksmatenFlexDiet, but not recall interviews, overreported energy intake compared with the estimated energy expenditure, but the limits of agreement were wider for the comparison with recall interviews than for RiksmatenFlexDiet. The ability to rank intake by estimated energy expenditure and biomarker concentrations was modest for both methods, but the correlations between intake and the reference methods were generally stronger for RiksmatenFlexDiet than for the recall interviews.

Reported energy intake was higher by RiksmatenFlexDiet than by recall interviews, and consequently, intake of macronutrients was also higher. In studies of adolescents, where energy intake by computer or Web-based 24-hour dietary recall methods have been compared with 24-hour dietary recall interviews, the difference between the methods have generally been small [15,44,45]. Limits of agreement for energy intake and macronutrients were wide in our study but similar to those observed in earlier studies of adolescents [44] or wider [15]. Agreement and ability to rank dietary intake were also better with these methods [15,44] compared with RiksmatenFlexDiet. However, in contrast to our study where dietary intake by the two methods was assessed on different days, the reference method in these studies covered the same days, and a lower agreement and ranking ability would therefore be expected when different days are compared. When a Web-based 24-hour dietary recall was compared with food diaries, thus covering different

time periods, limits of agreement, agreement, and ability to rank were comparable to the results of our study [46].

The agreement and ability to rank individuals by energy intake was better for RiksmatenFlexDiet than for the recall interviews. The Bland-Altman plots showed large variation at the individual level for both methods, as generally seen in validation studies of energy intake [47], and this was also expected in our study, as 2 days is not enough to estimate habitual energy intake. The limits of agreement, which were narrower for RiksmatenFlexDiet than for recall interviews, were in line with an earlier study of a Web-based food recall where energy intake was underestimated [21], but wider than that reported for a Web-based food diary in 10-year-old children [20]. The modest correlation between energy intake and estimated energy expenditure for RiksmatenFlexDiet was similar to other Web-based tools [20,21,26]. Although a somewhat stronger correlation between energy intake by two 24-hour dietary recall interviews has been reported earlier [26], this correlation was not significant for recall interviews in our study.

The evaluation of energy intake with estimated energy expenditure from accelerometer data should be interpreted with caution, since the energy expenditure estimation is not without errors. Furthermore, the comparison between energy intake and energy expenditure assumes energy balance, which may not be the case in growing adolescents. In addition, the equations used for estimating total energy expenditure were developed in a small group of 9-year-old children [40] and may be less accurate in adolescents [21]. Finally, activities less well captured by accelerometers, for example, biking, may be more common in adolescents, and some participants may also have taken off their accelerometers during intensive ball sports. The registered counts per minute were quite low in this study compared with European adolescents [48] and a subgroup from the Riksmaten adolescents 2016-2017 survey [49]. Thus, it is possible that energy expenditure is underestimated and that RiksmatenFlexDiet, as suggested by the comparison, does not overreport energy intake on an average.

The correlations between whole grain wheat and rye intake and alkylresorcinol homologs were acceptable for both methods and in line with other studies on children and adults [23,34,50,51]. Intake of bread and breakfast cereals, the major sources of whole grain wheat and rye in Swedish adolescents [36], is difficult to report. Photos of these foods are linked to the search engine in RiksmatenFlexDiet, to help the participants find the food closest to what they have eaten. This may have resulted in stronger correlations for RiksmatenFlexDiet than for recall interviews.

Reported intake of fruit and vegetables was similar by the two methods. The intraclass correlation between the two methods was stronger for fruit than for vegetables, in line with an earlier study on adolescents [15]. The correlations between intake of fruit and vegetables and the objective biomarkers plasma carotenoids were, however, weak for both methods. Moderate correlations between several carotenoids and various dietary assessment methods have been reported in adults [34,52-55], adolescent girls (but not boys) [56], and children and adolescents [22]. However, stronger correlations between intake from a Web-based food diary and carotenoids have also been reported in 10-year-old children [19]. Significantly stronger correlations in younger children (8-9 years old) compared to older children (12-14 years old) have also been reported [22]. The strongest correlations in our study were seen for lutein-zeaxanthin by both methods, with the strongest correlation for RiksmatenFlexDiet. This is in line with the correlations presented for adult women, but not men, where all carotenoid concentrations were significantly correlated with fruit and vegetable intake by a Web-based 4-day food diary [34]. Food records may also produce stronger correlations between intake and plasma concentrations than recall methods [32,46].

Strengths and Limitations

A major strength of this study is that dietary intake has been validated against objectively measured physical activity and biomarkers. Estimating energy expenditure from equations and accelerometer data is not without errors, as previously discussed, but the physical activity measurements are objectively measured. Biomarker concentrations provide a reflection of dietary intake, but perfect agreement could not be expected because genetic variability, lifestyle, and physiological factors influence the concentrations in plasma [27]. Furthermore, neither 2 days of dietary intake nor a single blood sample may reflect habitual intake of whole grain wheat and rye, fruits, and vegetables. Another possible limitation is that blood samples were not drawn in the fasting state. However, fasting was not a confounding factor in validation studies of carotenoid concentrations and dietary carotenoid intake [32], and correlations between alkylresorcinol concentrations and wholegrain wheat and rye intake were similar to those reported in other studies [23,34,50,51].

Another strength is that the comparison and validation of RiksmatenFlexDiet were performed in adolescents by using a recruitment procedure similar to the national dietary survey of adolescents that RiksmatenFlexDiet was intended for. Although this study was not designed to obtain a representative sample, the geographical spread across Sweden was good. Furthermore, around 50% of the participants came from homes where at least one parent had a postsecondary education. This proportion is lower than that in the subsequent national dietary survey in adolescents (60%), but higher than expected in households with children in the general population (38%) [36]. Thus, we believe that the results are generalizable to the population that the method is intended for. Moreover, we are not dependent on a convenience sample of self-selected and motivated participants that many comparison and validation studies rely on. However, the study is limited by the large proportion of participants who did not complete the random second registration day, making subanalyses by gender and age difficult. This day was lost for many participants, as they, particularly those in the youngest age group, did not read their email prompts and consequently missed the requirement of recording on their last day. In the main national dietary survey, the date of the randomly selected day was therefore visible when the participants logged on to RiksmatenFlex the first time. In addition, text message prompts and reminders were introduced. Thus, participants knew about the last day in advance, and the reminders also reached them. Knowing which day to record in advance could potentially

influence how the participants report their diet, but this limitation was considered less problematic for the final dietary information as compared to participants not completing their last day. With these changes, almost 90% of the participants completed the last random day in the main survey [36].

In general, Web-based methods are well accepted by participants [15,46]. How well computer-based methods work may depend on the structure of the method. Many foods and details may make a method difficult and laborious to complete. Some computer and Web-based methods include a large number of different foods [7,15], while others focus on fewer generic core foods [46]. A shorter list of generic foods, as in RiksmatenFlexDiet, makes the method more user-friendly, but as reported previously [46], some level of detail was inevitably lost. A future challenge is to develop Web-based methods that capture as much detail as possible without complicating completion too much.

Conclusions

This study demonstrates that RiksmatenFlex provides dietary intake estimates that are comparable with estimates from 24-hour dietary recall interviews in Swedish adolescents. The tool can easily be adapted to other age groups by changing food list and portion sizes. Thus, RiksmatenFlex has great potential for cost-effective dietary data collections in upcoming national dietary surveys and other studies in Sweden. Thus far, computer and Web-based tools do not provide dietary information with less bias than traditional dietary assessment methods; however, new technology has the potential to provide solutions that could further reduce dietary reporting bias. This is an important area for future research.

Acknowledgments

The authors thank project leader for Riksmaten Adolescents 2016-17 Ulla-Kaisa Koivisto Hursti, staff at the NFA and the OEMCs for collecting the data. We also thank schools and volunteers for taking part in the study.

The study was financed by the National Food Agency and the blood sample collection was made possible through funding from the Swedish Civil Contingency Agency.

Conflicts of Interest

RL (Chalmers University of Technology) holds research grants funded by Lantmännen Research Foundation and Barilla for specific projects. No other conflicts of interest are declared.

Multimedia Appendix 1
Photos of bread, ready-to-eat sandwiches, cereals, ice cream, and fat spreads attached to the search engine and food list.
[PDF File (Adobe PDF File), 597 KB - jmir_v21i10e12572_app1.pdf]

Multimedia Appendix 2
The food list in RiksmatenFlexDiet.
[PDF File (Adobe PDF File), 684 KB - jmir_v21i10e12572_app2.pdf]

Multimedia Appendix 3
Bland-Altman plots of reported intake of energy and macronutrients between the two dietary assessment methods.
[PDF File (Adobe PDF File), 179 KB - jmir_v21i10e12572_app3.pdf]

References

1. Schatzkin A, Kipnis V, Carroll RJ, Midthune D, Subar AF, Bingham S, et al. A comparison of a food frequency questionnaire with a 24-hour recall for use in an epidemiological cohort study: results from the biomarker-based Observing Protein and Energy Nutrition (OPEN) study. Int J Epidemiol 2003 Dec;32(6):1054-1062. [Medline: 14681273]
2. Subar AF, Kipnis V, Troiano RP, Midthune D, Schoeller DA, Bingham S, et al. Using intake biomarkers to evaluate the extent of dietary misreporting in a large sample of adults: the OPEN study. Am J Epidemiol 2003 Jul 01;158(1):1-13. [Medline: 12835280]
3. European Food Safety Authority. Guidance on the EU Menu methodology. EFS2 2014 Dec;12(12):3944. [doi: 10.2903/j.efsa.2014.3944]
4. Eldridge AL, Piernas C, Illner A, Gibney MJ, Gurinović MA, de Vries JHM, et al. Evaluation of New Technology-Based Tools for Dietary Intake Assessment-An ILSI Europe Dietary Intake and Exposure Task Force Evaluation. Nutrients 2018 Dec 28;11(1):55 [FREE Full text] [doi: 10.3390/nu11010055] [Medline: 30597864]
5. Illner A, Freisling H, Boeing H, Huybrechts I, Crispim SP, Slimani N. Review and evaluation of innovative technologies for measuring diet in nutritional epidemiology. Int J Epidemiol 2012 Aug;41(4):1187-1203 [FREE Full text] [doi: 10.1093/ije/dys105] [Medline: 22933652]

6. Timon CM, van den Barg R, Blain RJ, Kehoe L, Evans K, Walton J, et al. A review of the design and validation of web- and computer-based 24-h dietary recall tools. Nutr Res Rev 2016 Dec;29(2):268-280. [doi: 10.1017/S0954422416000172] [Medline: 27955721]
7. Arab L, Tseng C, Ang A, Jardack P. Validity of a multipass, web-based, 24-hour self-administered recall for assessment of total energy intake in blacks and whites. Am J Epidemiol 2011 Dec 01;174(11):1256-1265 [FREE Full text] [doi: 10.1093/aje/kwr224] [Medline: 22021561]
8. Nybacka S, Bertéus Forslund H, Wirfält E, Larsson I, Ericson U, Warensjö Lemming E, et al. Comparison of a web-based food record tool and a food-frequency questionnaire and objective validation using the doubly labelled water technique in a Swedish middle-aged population. J Nutr Sci 2016;5:e39 [FREE Full text] [doi: 10.1017/jns.2016.29] [Medline: 27752306]
9. Park Y, Dodd KW, Kipnis V, Thompson FE, Potischman N, Schoeller DA, et al. Comparison of self-reported dietary intakes from the Automated Self-Administered 24-h recall, 4-d food records, and food-frequency questionnaires against recovery biomarkers. Am J Clin Nutr 2018 Jan 01;107(1):80-93. [doi: 10.1093/ajcn/nqx002] [Medline: 29381789]
10. Yuan C, Spiegelman D, Rimm EB, Rosner BA, Stampfer MJ, Barnett JB, et al. Relative Validity of Nutrient Intakes Assessed by Questionnaire, 24-Hour Recalls, and Diet Records as Compared With Urinary Recovery and Plasma Concentration Biomarkers: Findings for Women. Am J Epidemiol 2018 May 01;187(5):1051-1063. [doi: 10.1093/aje/kwx328] [Medline: 29036411]
11. Wark PA, Hardie LJ, Frost GS, Alwan NA, Carter M, Elliott P, et al. Validity of an online 24-h recall tool (myfood24) for dietary assessment in population studies: comparison with biomarkers and standard interviews. BMC Med 2018 Aug 09;16(1):136 [FREE Full text] [doi: 10.1186/s12916-018-1113-8] [Medline: 30089491]
12. Baranowski T, Islam N, Baranowski J, Cullen KW, Myres D, Marsh T, et al. The food intake recording software system is valid among fourth-grade children. J Am Diet Assoc 2002 Mar;102(3):380-385. [Medline: 11902371]
13. Moore HJ, Ells LJ, McLure SA, Crooks S, Cumbor D, Summerbell CD, et al. The development and evaluation of a novel computer program to assess previous-day dietary and physical activity behaviours in school children: the Synchronised Nutrition and Activity Program (SNAP). Br J Nutr 2008 Jun;99(6):1266-1274. [doi: 10.1017/S0007114507862428] [Medline: 18042307]
14. Diep CS, Hingle M, Chen T, Dadabhoy HR, Beltran A, Baranowski J, et al. The Automated Self-Administered 24-Hour Dietary Recall for Children, 2012 Version, for Youth Aged 9 to 11 Years: A Validation Study. J Acad Nutr Diet 2015 Oct;115(10):1591-1598 [FREE Full text] [doi: 10.1016/j.jand.2015.02.021] [Medline: 25887784]
15. Albar SA, Alwan NA, Evans CEL, Greenwood DC, Cade JE. Agreement between an online dietary assessment tool (myfood24) and an interviewer-administered 24-h dietary recall in British adolescents aged 11-18 years. Br J Nutr 2016 Dec;115(9):1678-1686. [doi: 10.1017/S0007114516000593] [Medline: 26975650]
16. Bradley J, Simpson E, Poliakov I, Matthews JNS, Olivier P, Adamson AJ, et al. Comparison of INTAKE24 (an Online 24-h Dietary Recall Tool) with Interviewer-Led 24-h Recall in 11-24 Year-Old. Nutrients 2016 Jun 09;8(6):NA [FREE Full text] [doi: 10.3390/nu8060358] [Medline: 27294952]
17. Storey KE, McCargar LJ. Reliability and validity of Web-SPAN, a web-based method for assessing weight status, diet and physical activity in youth. J Hum Nutr Diet 2012 Feb;25(1):59-68. [doi: 10.1111/j.1365-277X.2011.01181.x] [Medline: 21615806]
18. Carvalho MA, Baranowski T, Foster E, Santos O, Cardoso B, Rito A, et al. Validation of the Portuguese self-administered computerised 24-hour dietary recall among second-, third- and fourth-grade children. J Hum Nutr Diet 2015 Dec;28(6):666-674. [doi: 10.1111/jhn.12280] [Medline: 25420921]
19. Biltoft-Jensen A, Bysted A, Trolle E, Christensen T, Knuthsen P, Damsgaard CT, et al. Evaluation of Web-based Dietary Assessment Software for Children: comparing reported fruit, juice and vegetable intakes with plasma carotenoid concentration and school lunch observations. Br J Nutr 2013 Jul 14;110(1):186-195. [doi: 10.1017/S0007114512004746] [Medline: 23181984]
20. Biltoft-Jensen A, Hjorth MF, Trolle E, Christensen T, Brockhoff PB, Andersen LF, et al. Comparison of estimated energy intake using Web-based Dietary Assessment Software with accelerometer-determined energy expenditure in children. Food Nutr Res 2013;57:NA [FREE Full text] [doi: 10.3402/fnr.v57i0.21434] [Medline: 24358037]
21. Medin AC, Hansen BH, Astrup H, Ekelund U, Frost Andersen L. Validation of energy intake from a web-based food recall for children and adolescents. PLoS One 2017;12(6):e0178921 [FREE Full text] [doi: 10.1371/journal.pone.0178921] [Medline: 28594899]
22. Medin AC, Carlsen MH, Andersen LF. Associations between reported intakes of carotenoid-rich foods and concentrations of carotenoids in plasma: a validation study of a web-based food recall for children and adolescents. Public Health Nutr 2016 Dec;19(18):3265-3275. [doi: 10.1017/S1368980016001622] [Medline: 27346293]
23. Biltoft-Jensen A, Damsgaard CT, Andersen EW, Ygil KH, Andersen R, Ege M, et al. Validation of Reported Whole-Grain Intake from a Web-Based Dietary Record against Plasma Alkylresorcinol Concentrations in 8- to 11-Year-Olds Participating in a Randomized Controlled Trial. J Nutr 2016 Feb;146(2):377-383. [doi: 10.3945/jn.115.222620] [Medline: 26764319]
24. Gibson RS. Principles of Nutritional Assessment. New York: Oxford University Press; 2005.
25. Andersen LF, Pollestad ML, Jacobs DR, Løvø A, Hustvedt B. Validation of a pre-coded food diary used among 13-year-olds: comparison of energy intake with energy expenditure. Public Health Nutr 2005 Dec;8(8):1315-1321. [Medline: 16372928]

26. Rothausen BW, Matthiessen J, Groth MV, Brockhoff PB, Andersen LF, Trolle E. Comparison of estimated energy intake from 2×24-hour recalls and a seven-day food record with objective measurements of energy expenditure in children. Food Nutr Res 2012;56:NA [FREE Full text] [doi: 10.3402/fnr.v56i0.12221] [Medline: 22347842]
27. Jenab M, Slimani N, Bictash M, Ferrari P, Bingham SA. Biomarkers in nutritional epidemiology: applications, needs and new horizons. Hum Genet 2009 Jun;125(5-6):507-525. [doi: 10.1007/s00439-009-0662-5] [Medline: 19357868]
28. Zhang B, Zhao Q, Guo W, Bao W, Wang X. Association of whole grain intake with all-cause, cardiovascular, and cancer mortality: a systematic review and dose-response meta-analysis from prospective cohort studies. Eur J Clin Nutr 2018 Dec;72(1):57-65. [doi: 10.1038/ejcn.2017.149] [Medline: 29091078]
29. Ross AB. Present status and perspectives on the use of alkylresorcinols as biomarkers of wholegrain wheat and rye intake. J Nutr Metab 2012;2012:462967 [FREE Full text] [doi: 10.1155/2012/462967] [Medline: 22363838]
30. Kyrø C, Olsen A, Bueno-de-Mesquita HBA, Skeie G, Loft S, Åman P, et al. Plasma alkylresorcinol concentrations, biomarkers of whole-grain wheat and rye intake, in the European Prospective Investigation into Cancer and Nutrition (EPIC) cohort. Br J Nutr 2014 May 28;111(10):1881-1890. [doi: 10.1017/S0007114513004388] [Medline: 24521535]
31. Biskup I, Kyrø C, Marklund M, Olsen A, van Dam RM, Tjønneland A, et al. Plasma alkylresorcinols, biomarkers of whole-grain wheat and rye intake, and risk of type 2 diabetes in Scandinavian men and women. Am J Clin Nutr 2016 Dec;104(1):88-96. [doi: 10.3945/ajcn.116.133496] [Medline: 27281306]
32. Burrows TL, Williams R, Rollo M, Wood L, Garg ML, Jensen M, et al. Plasma carotenoid levels as biomarkers of dietary carotenoid consumption: A systematic review of the validation studies. Journal of Nutrition & Intermediary Metabolism 2015 Aug;2(1-2):15-64. [doi: 10.1016/j.jnim.2015.05.001]
33. Burrows TL, Rollo ME, Williams R, Wood LG, Garg ML, Jensen M, et al. A Systematic Review of Technology-Based Dietary Intake Assessment Validation Studies That Include Carotenoid Biomarkers. Nutrients 2017 Feb 14;9(2):NA [FREE Full text] [doi: 10.3390/nu9020140] [Medline: 28216582]
34. Nybacka S, Lindroos AK, Wirfält E, Leanderson P, Landberg R, Ericson U, et al. Carotenoids and alkylresorcinols as objective biomarkers of diet quality when assessing the validity of a web-based food record tool and a food frequency questionnaire in a middle-aged population. BMC Nutr 2016;2:53. [doi: 10.1186/s40795-016-0094-2]
35. Warensjö Lemming E, Nälsén C, Becker W, Ridefelt P, Mattisson I, Lindroos AK. Relative validation of the dietary intake of fatty acids among adults in the Swedish National Dietary Survey using plasma phospholipid fatty acid composition. J Nutr Sci 2015;4:e25 [FREE Full text] [doi: 10.1017/jns.2015.1] [Medline: 26495117]
36. Moraeus L, Lemming EW, Hursti UK, Arnemo M, Sipinen JP, Lindroos A. Riksmaten Adolescents 2016-17: A national dietary survey in Sweden - design, methods, and participation. Food Nutr Res 2018;62:NA. [doi: 10.29219/fnr.v62.1381] [Medline: 30186087]
37. Amcoff E, Edberg A, Enghardt Barbieri H, Lindroos AK, Nälsén C, Pearson M, et al. Riksmaten - vuxna 2010-11: Livsmedels- och näringsintag bland vuxna i Sverige. Uppsala: Livsmedelsverket; 2012. URL: https://www.livsmedelsverket.se/globalassets/publikationsdatabas/rapporter/2011/riksmaten_2010_20111.pdf [accessed 2019-09-03]
38. Serra-Majem L, Frost Andersen L, Henríque-Sánchez P, Doreste-Alonso J, Sánchez-Villegas A, Ortiz-Andrelluchi A, et al. Evaluating the quality of dietary intake validation studies. Br J Nutr 2009 Dec;102 Suppl 1:S3-S9. [doi: 10.1017/S0007114509993114] [Medline: 20100366]
39. Cole TJ, Lobstein T. Extended international (IOTF) body mass index cut-offs for thinness, overweight and obesity. Pediatr Obes 2012 Aug;7(4):284-294. [doi: 10.1111/j.2047-6310.2012.00064.x] [Medline: 22715120]
40. Nilsson A, Brage S, Riddoch C, Anderssen SA, Sardinha LB, Wedderkopp N, et al. Comparison of equations for predicting energy expenditure from accelerometer counts in children. Scand J Med Sci Sports 2008 Oct;18(5):643-650. [doi: 10.1111/j.1600-0838.2007.00694.x] [Medline: 18208433]
41. Maffeis C, Schutz Y, Zoccante L, Micciolo R, Pinelli L. Meal-induced thermogenesis in lean and obese prepubertal children. Am J Clin Nutr 1993 Apr;57(4):481-485. [doi: 10.1093/ajcn/57.4.481] [Medline: 8460601]
42. Henry CJK. Basal metabolic rate studies in humans: measurement and development of new equations. Public Health Nutr 2005 Oct;8(7A):1133-1152. [Medline: 16277825]
43. Wierzbicka R, Wu H, Franek M, Kamal-Eldin A, Landberg R. Determination of alkylresorcinols and their metabolites in biological samples by gas chromatography-mass spectrometry. J Chromatogr B Analyt Technol Biomed Life Sci 2015 Sep 01;1000:120-129 [FREE Full text] [doi: 10.1016/j.jchromb.2015.07.009] [Medline: 26218771]
44. Foster E, Delve J, Simpson E, Breininger S. Comparison study: INTAKE24 vs Interviewer led recall. URL: https://intake24.co.uk/assets/papers/Intake24-Comparison-report.pdf [accessed 2018-10-18] [WebCite Cache ID 73FwrsPcz]
45. Vereecken CA, Covents M, Matthys C, Maes L. Young adolescents' nutrition assessment on computer (YANA-C). Eur J Clin Nutr 2005 May;59(5):658-667. [doi: 10.1038/sj.ejcn.1602124] [Medline: 15741983]
46. Timon CM, Blain RJ, McNulty B, Kehoe L, Evans K, Walton J, et al. The Development, Validation, and User Evaluation of Foodbook24: A Web-Based Dietary Assessment Tool Developed for the Irish Adult Population. J Med Internet Res 2017 May 11;19(5):e158 [FREE Full text] [doi: 10.2196/jmir.6407] [Medline: 28495662]
47. Livingstone MBE, Robson PJ, Wallace JMW. Issues in dietary intake assessment of children and adolescents. Br J Nutr 2004 Oct;92 Suppl 2:S213-S222. [Medline: 15522159]

48. Jiménez-Pavón D, Fernández-Vázquez A, Alexy U, Pedrero R, Cuenca-García M, Polito A, HELENA Study Group. Association of objectively measured physical activity with body components in European adolescents. BMC Public Health 2013 Jul 18;13:667 [FREE Full text] [doi: 10.1186/1471-2458-13-667] [Medline: 23866681]
49. Nyberg G. De aktiva och De inaktiva: Om ungas rörelse i skola och på fritid. 2017:2.: Centrum för idrottsforskning Få unga rör sig tillräckligt URL: https://centrumforidrottsforskning.se/wp-content/uploads/2017/06/De-aktiva-och-de-inaktiva_komp.pdf [accessed 2019-09-03]
50. Landberg R, Kamal-Eldin A, Andersson A, Vessby B, Aman P. Alkylresorcinols as biomarkers of whole-grain wheat and rye intake: plasma concentration and intake estimated from dietary records. Am J Clin Nutr 2008 Apr;87(4):832-838. [doi: 10.1093/ajcn/87.4.832] [Medline: 18400704]
51. Landberg R, Kamal-Eldin A, Aman P, Christensen J, Overvad K, Tjønneland A, et al. Determinants of plasma alkylresorcinol concentration in Danish post-menopausal women. Eur J Clin Nutr 2011 Jan;65(1):94-101. [doi: 10.1038/ejcn.2010.193] [Medline: 20859297]
52. Jansen MCJF, Van Kappel AL, Ocké MC, Van 't Veer P, Boshuizen HC, Riboli E, et al. Plasma carotenoid levels in Dutch men and women, and the relation with vegetable and fruit consumption. Eur J Clin Nutr 2004 Oct;58(10):1386-1395. [doi: 10.1038/sj.ejcn.1601981] [Medline: 15054421]
53. Al-Delaimy WK, Ferrari P, Slimani N, Pala V, Johansson I, Nilsson S, et al. Plasma carotenoids as biomarkers of intake of fruits and vegetables: individual-level correlations in the European Prospective Investigation into Cancer and Nutrition (EPIC). Eur J Clin Nutr 2005 Dec;59(12):1387-1396. [doi: 10.1038/sj.ejcn.1602252] [Medline: 16160702]
54. Lassale C, Castetbon K, Laporte F, Deschamps V, Vernay M, Camilleri GM, et al. Correlations between Fruit, Vegetables, Fish, Vitamins, and Fatty Acids Estimated by Web-Based Nonconsecutive Dietary Records and Respective Biomarkers of Nutritional Status. J Acad Nutr Diet 2016 Mar;116(3):427-438.e5. [doi: 10.1016/j.jand.2015.09.017] [Medline: 26522988]
55. Andersen LF, Veierød MB, Johansson L, Sakhi A, Solvoll K, Drevon CA. Evaluation of three dietary assessment methods and serum biomarkers as measures of fruit and vegetable intake, using the method of triads. Br J Nutr 2005 Apr;93(4):519-527. [Medline: 15946415]
56. Vandevijvere S, Geelen A, Gonzalez-Gross M, van't Veer P, Dallongeville J, Mouratidou T, et al. Evaluation of food and nutrient intake assessment using concentration biomarkers in European adolescents from the Healthy Lifestyle in Europe by Nutrition in Adolescence study. Br J Nutr 2013 Feb 28;109(4):736-747. [doi: 10.1017/S0007114512002012] [Medline: 22617187]

Abbreviations

Cpm: counts per minute
EEest: estimated total energy expenditure
EFSA: European Food Safety Authority
NFA: National Food Agency, Sweden
OEMC: Occupational and Environmental Medicine Centers

Edited by G Eysenbach; submitted 28.10.18; peer-reviewed by A Laine, S Christensen, A Moen, A Medin; comments to author 21.03.19; revised version received 14.06.19; accepted 19.08.19; published 04.10.19.

Please cite as:
Lindroos AK, Petrelius Sipinen J, Axelsson C, Nyberg G, Landberg R, Leanderson P, Arnemo M, Warensjö Lemming E
Use of a Web-Based Dietary Assessment Tool (RiksmatenFlex) in Swedish Adolescents: Comparison and Validation Study
J Med Internet Res 2019;21(10):e12572
URL: https://www.jmir.org/2019/10/e12572
doi:10.2196/12572
PMID:31588902

Original Paper

Identifying the Most Autonomy-Supportive Message Frame in Digital Health Communication: A 2x2 Between-Subjects Experiment

Eline Suzanne Smit[1], PhD; Chamoetal Zeidler[1], MSc; Ken Resnicow[2], PhD; Hein de Vries[3], PhD

[1]Department of Communication Science, Amsterdam School of Communication Research, University of Amsterdam, Amsterdam, Netherlands
[2]School of Public Health, University of Michigan, Ann Arbor, MI, United States
[3]Department of Health Promotion, Care and Public Health Research Institute, Maastricht University, Maastricht, Netherlands

Corresponding Author:
Eline Suzanne Smit, PhD
Department of Communication Science
Amsterdam School of Communication Research
University of Amsterdam
PO Box 15791
Amsterdam, 1001 NG
Netherlands
Phone: 31 (0)625258558
Email: E.S.Smit@uva.nl

Abstract

Background: The effectiveness of digital health communication may be increased by enhancing autonomy supportiveness.

Objective: This study aimed to identify the most autonomy-supportive message frame within an intervention for increasing vegetable intake by testing the effect of the following 2 strategies: (1) using autonomy-supportive language and (2) providing choice.

Methods: A Web-based 2 (autonomy-supportive vs controlling language)×2 (choice vs no choice) experiment was conducted among 526 participants, recruited via a research panel. The main outcome measures were perceived autonomy support (measured using the Virtual Care Climate Questionnaire, answered with scores 1 to 5), perceived relevance (measured with one question, answered with scores 1 to 5), and overall evaluation of the intervention (measured with 1 open-ended question, answered with scores 1 to 10).

Results: Choice had a significant positive effect on the overall evaluation of the intervention (b=.12; P=.003), whereas for participants with a high need for autonomy, there was a significant positive effect on perceived relevance (b=.13; P=.02). The positive effect of choice on perceived autonomy support approached significance (b=.07; P=.07). No significant effects on any of the three outcomes were observed for language.

Conclusions: Results suggest that provision of choice rather than the use of autonomy-supportive language can be an easy-to-implement strategy to increase the effectiveness of digital forms of health communication, especially for people with a high need for autonomy.

(J Med Internet Res 2019;21(10):e14074) doi:10.2196/14074

KEYWORDS

health communication; health behavior; personal autonomy; internet; health promotion; healthy diet; self-determination theory

Introduction

Digital Health Communication to Date

Digital forms of health communication, for example, Web-based computer-tailored interventions, can be a cost-effective strategy for health promotion [1-3]. However, the effect sizes found in previous studies have so far remained small [1], which suggests room for improvement. As the public health impact of digital health communication can only be maximized when we use the immense reach of the internet and optimize its efficacy (impact=reach×efficacy [4]), testing strategies that might increase efficacy is a priority.

Such effect improvement may be achieved by moving beyond a focus on *what* health information is provided to a focus on *how* this information is provided. Until recently, health communication scholars have mainly focused on tailoring the content of digital health communication based on receivers' current health behavior and behavioral determinants [5]. However, variations in the communication style (ie, the *how*) that is used to deliver health communication messages are believed to have different effects as well [6]. Previous studies indicate, for instance, that autonomy-supportive communication strategies may enhance the impact of face-to-face health communication interventions, such as counseling, by facilitating the internalization of motivation [7]. Yet, little is known about the effect of autonomy-supportive strategies in digital forms of health communication [8]. Therefore, this study aimed to identify the most autonomy-supportive message frame by testing the effects of the following 2 strategies intended to increase perceptions of autonomy support: (1) the use of autonomy-supportive language; and (2) the provision of choice, within a Web-based computer-tailored intervention aimed at increasing vegetable intake.

Autonomy-Supportive Communication Styles

Previously, both autonomy-supportive and more directive communication styles have been identified and studied in offline forms of health communication [9,10]. According to Self-Determination Theory (SDT) [8,11], supporting autonomy is an important prerequisite for achieving autonomous motivation for health behavior change. Autonomous motivation has been found to be an important predictor of actual behavior change and subsequent positive health outcomes [7,12]. In the face-to-face setting, providing autonomy support involves strategies such as the elicitation and acknowledgment of a person's perspectives, the provision of a clear rationale for change, supporting the person's volition, and using autonomy-supportive or noncontrolling language and offering choice [13]. Such strategies have, however, not been studied to a large extent within the context of digital health communication. We found 2 digital health studies that manipulated autonomy support. In the first, the authors report positive effects of the provision of autonomy support by a virtual health care provider. However, the authors do not provide much detail regarding how the support was provided [11]. The second study elaborately describes the operationalization of autonomy support provided by a computer-based personal trainer using choice (eg, "choose whichever one works better for you" vs "do this"), the acknowledgement of feelings (eg, "some people feel intimidated and those feelings are normal" vs "some people feel intimidated, but that's not useful"), and minimal evaluations or judgments (ie, providing recommendations vs telling participants what they should do), although without reporting on the results of the study [8]. A recent meta-analysis of techniques to promote motivation for health behavior change even distinguished 18 different techniques aimed at the promotion of need satisfaction and autonomous motivation, the provision of choice and the use of noncontrolling language being 2 of them [14]. We are aware of only 1 digital health study that focused on the effects of different autonomy-supportive strategies in virtual care settings without virtual health care providers involved. At the same time, most Web-based computer-tailored health communication interventions do not involve a virtual health care provider, but rather provide tailored advice with the program as a main source. In that study [15], the authors report no differential effects of an autonomy-supportive versus controlling tone when used in Web-based computer-tailored alcohol reduction messages on perceived autonomy support from and reactance toward these messages, neither did they find a moderation effect for baseline need for autonomy. However, it should be noted that participants in the study, regardless of the message tone used, generally rated the messages positively, providing high levels of autonomy support. This suggests that there may not have been adequate conceptual separation between message conditions [15].

Providing Autonomy Support in Web-Based Computer Tailoring

Computer-tailored health communication uses a computerized process to adjust message content based on the individual users' personal characteristics (eg, behavior, personality, attitudes, and beliefs), with the goal of increasing perceptions of personal relevance [16]. As such, tailored health communication interventions have been found to be more successful in attracting and keeping the receiver's attention [16,17] and in achieving effortful processing of information [18] when compared with nontailored interventions. When tailored content is provided using an autonomy-supportive communication style, similar to face-to-face autonomy-supportive communication, receivers ideally perceive the intervention as more supportive of their autonomy, and co-occurring feelings of personal freedom may rise. This increased sense of freedom might encourage the receiver to process in-depth only those parts of the intervention that are most appealing to himself or herself personally, thereby further increasing perceptions of personal relevance.

Given the positive effects found of an autonomy-supportive communication style (ie, including the use of autonomy-supportive language and the provision of choice) in the face-to-face setting [9,10] and virtual clinician context [8,11], in addition to theory and evidence on tailoring effects, we formulated the first 2 hypotheses.

- H1: The use of autonomy-supportive language will result in (a) higher perceived autonomy support, (b) higher perceived relevance, and (c) a more positive overall evaluation of the intervention compared with the use of controlling language.
- H2: The provision of choice will result in (a) higher perceived autonomy support, (b) higher perceived relevance, and (c) a more positive overall evaluation of the intervention compared with no provision of choice.

The Need for Autonomy as a Moderator

Although SDT suggests a universal need for autonomy, there may be individual differences in how autonomy needs influence message impact [19]. Some people prefer to choose their own path toward lifestyle improvement, whereas others prefer to be guided by clear-cut expert advice [20]. Although SDT scholars have recognized these individual differences [7,12], only few have taken these differences into account when developing

health communication strategies, and none have done so in the context of digital health communication. Not taking into account individual differences in the need for autonomy may result in communicating digital health information that does not fit people's personal needs, making it less likely to be read and less likely to be considered as personally relevant [21]. As a consequence, messages are less likely to be centrally processed, reducing behavioral impact. This idea finds support in findings from 2 studies that investigated the effects of printed health communication aimed to increase colorectal cancer screening [20] and fruit and vegetable intake [22]. Both studies found that for people with greater need for autonomy—operationalized as a preference for autonomy-supportive communication compared with directive communication—newsletters that were communicated in an autonomy-supportive tone were more effective in changing target behaviors than newsletters that were communicated in a more directive tone.

On the basis of these previous studies, we propose that in the context of digital health communication, baseline need for autonomy will interact with the intervention manipulations as follows:

- H3: The effects of the use of autonomy-supportive language (vs the use of controlling language) on (a) perceived autonomy support, (b) perceived relevance, and (c) the overall evaluation of the intervention will be stronger for respondents with a high need for autonomy than for respondents with a low need for autonomy.
- H4: The effects of the provision of choice (vs no provision of choice) on (a) perceived autonomy support, (b) perceived relevance, and (c) the overall evaluation of the intervention will be stronger for respondents with a high need for autonomy than for respondents with a low need for autonomy.

Methods

Design and Participants

To test the hypotheses, an experiment with a 2 (language use: autonomy-supportive vs controlling language)×2 (choice: provided vs not provided) between-subjects design was conducted within the context of an existing Web-based computer-tailored intervention module aimed at increasing vegetable consumption [23]. Participants were recruited via *PanelClix*, an International Organization for Standardization–certified research panel [24].

A total of 728 Dutch adult participants started the experiment. Participants who did not give their informed consent (1/728, 0.1%) were not interested in eating—or continuing to eat—250 g of vegetables per day or did not provide an answer to this inclusion question (7/728, 1.0%) were excluded and not randomized. Of 720 randomized participants, 604 (83.8%) completed the entire intervention and questionnaire. However, we excluded participants with problematic or implausible response patterns: 7/604 (0.9%) participants took too long to fill in the questionnaire (3 SDs or more above the mean completion time), 23/604 (3.0%) participants filled in the questionnaire too fast (ie, <5 min), 2/604 (0.3%) participants filled in 0 as their weight, 9/604 (13.0%) participants had an extremely high vegetable consumption (3 SDs or more above the mean vegetable consumption), and 46/604 (6.0%) participants did not answer the 7 process evaluation questions in a logically consistent manner (ie, they filled in the same response for all questions even when items were scaled in opposing directions). For some of the participants, more than one of these problems were encountered; a total of 79/604 (13.1%) participants were excluded, and the final sample consisted of 525 participants. Of them, 231/525 (44.0%) were men, and 294/525 (56.0%) were women, with age ranging from 18 to 65 years (mean 43.35 years, SD 13.80). About half of the participants (261/525, 49.7%) were highly educated, and participants consumed on an average 183.80 g of vegetables daily (SD 96.45). Their average body mass index (BMI) was 25.52 kg/m^2 (SD 4.95).

A Consolidated Standards of Reporting Trials flow diagram is provided in Multimedia Appendix 1, whereas Table 1 provides an overview of the final sample's characteristics.

Table 1. Sample characteristics (N=525).

Variables	Value, n (%)	Value, mean (SD)	Range
Sex			
Male	231 (44.0)	—[a]	—
Female	294 (56.0)	—	—
Age (years)	—	43.35 (13.80)	18-65
Length (cm)	—	174.64 (9.69)	154-200
Weight (kg)	—	78.07 (17.63)	30-192
Body mass index (kg/m^2)	—	25.52 (4.95)	12.49-52.08
Education			
Low	41 (7.8)	—	—
Middle	222 (42.3)	—	—
High	261 (49.7)	—	—
Other	1 (0.1)	—	—
Vegetable consumption	—	183.80 (96.45)	0-700

[a]Not applicable.

Procedure

This study was approved by the ethics committee of the University of Amsterdam (reference number: 2016-PC-7205). First, participants received a brief explanation about the study aims and procedures as well as information about their rights and the confidential handling of their data. After participants provided their Web-based informed consent, they were asked about their intention to (continue) eating 250 g of vegetables per day, that is, the Dutch guideline for vegetable consumption [25]. When participants had a positive intention (yes vs no), they were randomly assigned to 1 of the following 4 conditions: autonomy-supportive language, or controlling language, with provision of choice, or without provision of choice (see Table 2 for the number of participants per condition). Subsequently, participants were asked about their demographics. After completion of the intervention (which is described in more detail in the following section), participants reported their perceived autonomy support and personal as well as their overall evaluation of the intervention. The postintervention questionnaire also queried participants' need for autonomy. The average time that participants needed to complete the questionnaire was approximately 15 min (to be precise: 964.69 seconds; SD 981.74). Participants were rewarded by means of 150 Clix, the incentive of PanelClix, worth approximately €1.88.

Table 2. Experimental conditions.

Choice	Language use, n (%)		Total, n (%)
	Autonomy supportive	Controlling	
Yes	147 (28.0)	124 (26.6)	271 (51.6)
No	141 (26.9)	113 (21.5)	254 (48.4)
Total	288 (54.9)	237 (45.1)	525 (100.00)

The Web-Based Computer-Tailored Intervention

The Web-based computer-tailored intervention was based on an intervention previously developed by Schulz et al, which aimed to improve several lifestyle-associated behaviors (ie, physical activity, vegetable consumption, fruit consumption, alcohol intake, and smoking cessation) and was called myHealthyBehaviour. As the study into the effectiveness of myHealthyBehaviour showed that the percentage of noncompliance with Dutch health guidelines was highest for vegetable intake (ie, 68%) [23], this study specifically focused on the vegetable consumption module of the intervention. For this study, we therefore changed the intervention's name to MyVegetableConsumption. In addition, we incorporated the updated Dutch guideline of consuming a minimum of 250 g rather than 200 g of vegetables daily [25].

The intervention consisted of 4 steps, each consisting of a set of questions and tailored feedback based on their answers, following an initial assessment of respondents' vegetable consumption. The first step looked at the advantages and disadvantages participants experienced with regard to consuming enough vegetables, for example, the (expensive) price of vegetables and their positive effect on one's health. The second step looked at the influence of the social environment of the participants, including one's partner, family, friends, and colleagues. The third step assisted participants in making preparatory plans to consume enough vegetables, for example, by bringing vegetables to work as part of their lunch. The fourth

step looked at participants' self-efficacy and aided them in making plans to cope with potentially difficult situations, for example, in busy times.

The Manipulations

The tailored feedback participants received was written in either an autonomy-supportive or controlling language, and choice options were either provided or not provided throughout the intervention.

Language

Language was manipulated into autonomy-supportive language or controlling language that was used throughout the intervention, that is, in (the introductions to) the questions and the tailored feedback. Autonomy-supportive language was intended to provide participants a sense of volition over their decisions and used more tentative advice, for example, "you could try to bring snack vegetables to work." In contrast, controlling language was more directive and definitive, for example, "you must bring snack vegetables to work!" [9,26].

Provision of Choice

Provision of choice was manipulated by either providing or not providing participants the possibility to choose if they wanted to make each out of 5 suggested preparatory plans (step 3) and to choose whether or not they wanted to make coping plans for each of the 7 potentially difficult situations described (step 4). Accordingly, participants who were provided with choice received feedback that was tailored based on the plans they chose (not) to make. Participants who were not provided with choice received 1 (nontailored) advice statement for preparatory planning and 1 (nontailored) advice statement addressing plans to cope with potentially difficult situations.

Pilot Test

A pilot test of the manipulations was conducted among experts in digital health communication and health message framing (N=8) and target group members (N=8), the latter varying in age, sex, and socioeconomic status. Experts were identified through the professional network of the first author, and target group members were identified through the first and second authors' private networks. Both were invited to complete the intervention and accompanying evaluation questions and were asked to identify any ambiguities in the questions and/or feedback. Moreover, they were asked to indicate whether the intervention felt autonomy supportive or controlling and whether they experienced choice or not. On the basis of the results from the pilot test, several improvements were made to messages and assessments.

An example of a feedback message in autonomy-supportive and controlling language, combined with and without choice, is provided in Multimedia Appendix 2.

Variables and Measures

Demographic and Other Background Variables

Several demographic variables were assessed, that is, sex, age, educational level, and marital status. BMI was estimated based on the participant's self-reported height and weight, and weekly vegetable consumption was measured using a 4-item food frequency questionnaire [27].

Perceived Autonomy Support

Perceived autonomy support was the first dependent variable and was measured using the Virtual Care Climate Questionnaire [28]. The 15 items (eg, "I feel that MyVegetableConsumption has provided me with choices and options") could be answered from 1 (*completely disagree*) to 5 (*completely agree*) and were transformed into a mean score (alpha=.96; mean 3.76, SD 0.77).

Perceived Relevance

Perceived relevance was the second dependent variable and was measured with 1 question ("I perceived the feedback messages as personally relevant") that could be answered from 1 (*completely disagree*) to 5 (*completely agree*; mean 3.74, SD 1.02).

Overall Evaluation of the Intervention

Overall evaluation of the intervention was the third dependent variable and was measured by an open-ended question to give an overall grade for the intervention: "Please evaluate the intervention with a school grade from 1 to 10" (1=lowest grade and 10=highest grade; mean 7.50, SD 1.20).

Need for Autonomy

Need for autonomy, the main moderator variable, was measured with a 9-item scale. First, 6 items from the Health Causality Orientation Scale (HCOS) were included. The HCOS was developed by the third author based on the General Causality Orientation Scale [12] and looks at 4 causality orientations individuals may have when it concerns their health, namely, autonomous, controlled (experts), controlled (peers), and impersonal. To assess these orientations, participants were presented with 2 vignettes, in which they had to imagine that (1) they were discussing with a health professional how best to obtain their health-related goals and (2) they wanted to change their health behavior. Each vignette was followed by 4 items representing participants' autonomous, controlled (experts), controlled (peers), and impersonal orientations, on a scale from 1 (*very unlikely*) to 5 (*very likely*). For this study, only the 6 items measuring participants' autonomous and controlled orientation were considered, as the 2 items measuring impersonal orientations were not considered relevant for this study's purposes. Second, 3 items were included based on previous measures for assessing communication style preferences [20,22], for example, "When it comes to my health, I would like to have an expert tell me what to do" (1=*completely disagree* and 5=*completely agree*). A factor analysis with all 9 items revealed 2 factors: (1) a factor with 3 items, representing the need for autonomy (alpha=.61; mean 4.16, SD 0.66) and (2) a factor with 6 items, representing the need for external control (alpha=.76; mean 2.94, SD 0.77).

Data Analysis

First, we checked whether there was an equal distribution of demographic and other background variables across the conditions by conducting Chi-square tests and analyses of variance. If baseline differences were detected, correlations between these variables and the dependent variables were

calculated. Variables that were not equally distributed across conditions and were correlated significantly with one or more of the dependent variables were included in subsequent analyses as covariates.

Second, regression analyses were conducted to test the effects of language and choice on each of the 3 dependent variables (ie, perceived autonomy support, perceived relevance, and overall evaluation of the intervention). When there appeared to be a significant interaction effect between (either of) the 2 conditions and (one of) the moderator(s), the interaction was dismantled by conducting a median split of the moderator and comparing outcomes between the 2 groups.

Results

Correlations

Sex (χ^2_3=6.5 *P*=.09), educational level (χ^2_6=10.8; *P*=.10), marital status (χ^2_{12}=19.1; *P*=.09), age ($F_{3,521}$=0.54; *P*=.65), BMI ($F_{3,521}$=0.63; *P*=.59), and vegetable consumption ($F_{3,521}$=1.30; *P*=.27) were all distributed equally across the 4 conditions. As all variables were equally distributed across conditions, correlations between these variables and the dependent variables were not reported, and none of these variables were included in subsequent analyses as covariates. Mean scores including their SDs for each of the 3 dependent variables are, however, reported per condition in Table 3.

Table 3. Means (SDs) for dependent variables per condition (N=525).

Variables	Autonomy-supportive language and choice (n=147)	Autonomy-supportive language and no choice (n=141)	Controlling language and choice (n=124)	Controlling language and no choice (n=113)
Perceived autonomy support	3.77 (0.76)	3.70 (0.84)	3.82 (0.73)	3.76 (0.72)
Perceived relevance	3.77 (0.94)	3.67 (1.12)	3.80 (1.00)	3.69 (1.03)
Overall evaluation	7.60 (1.05)	7.33 (1.40)	7.66 (1.07)	7.42 (1.23)

Hypothesis Testing

Perceived Autonomy Support

The positive effect of choice on perceived autonomy support approached significance (*b*=.07; *P*=.07), providing partial support for hypothesis 2a. For language, no significant main effect was found on perceived autonomy support. Neither were any interaction effects found between choice and language and the need for autonomy or need for external control. Therefore, hypotheses 1a, 3a, and 4a should be rejected. There were, however, significant main effects of the need for autonomy (*b*=.32; *P*<.001) and the need for external control (*b*=.32; *P*<.001) on perceived autonomy support. This means that participants with a high need for autonomy and participants with a high need for external control reported high levels of perceived autonomy support, independent of the message received. Table 4 provides the results of this analysis.

Table 4. Effect of choice, language use, and need for autonomy on perceived autonomy support.

Variables	*Beta (B)*	SE (*B*)	*Beta (b)*	*t test (df)*	*P* value	95% CI
Choice	.06	0.03	.07	1.85 (11,513)	.07	0.85 to 1.73
Language	−.03	0.03	−.04	−0.97 (11,513)	.33	−0.09 to 0.03
Need for autonomy	.37	0.05	.32	7.95 (11,513)	<.001	0.28 to 0.46
Need for external control	.32	0.04	.32	8.04 (11,513)	<.001	0.24 to 0.39
Choice×language	−.01	0.03	−.01	−0.30 (11,513)	.76	−0.07 to 0.05
Choice×need for autonomy	−.03	0.03	−.04	−1.03 (11,513)	.30	−0.09 to 0.03
Choice×need for external control	−.03	0.03	−.04	−0.96 (11,513)	.34	−0.09 to 0.03
Language×need for autonomy	−.03	0.03	−.04	−1.09(11,513)	.28	−0.09 to 0.03
Language×need for external control	−.01	0.03	−.01	−0.15 (11,513)	.88	−0.07 to 0.06
Choice×language×need for autonomy	.01	0.03	.02	0.40 (11,513)	.69	−0.05 to 0.07
Choice×language×need for external control	−.02	0.03	−.02	−0.55 (11,513)	.58	−0.08 to 0.04

Perceived Relevance

There was a significant interaction effect between choice and the need for autonomy on perceived relevance (*b*=.08; *P*=.04; data not reported). Therefore, we report outcomes separately for participants with a high need (ie, with a score <4.32) and participants with a low need for autonomy (ie, with a score >4.32), based on a median split procedure. This showed that for participants with a low need for autonomy, there was no effect of choice on perceived relevance, whereas for participants with a high need for autonomy, there was a significant positive effect of choice on perceived relevance (*b*=.13; *P*=.02). Although hypotheses 1b, 2b, and 3b need to be rejected, the results thus confirm hypothesis 4b. For participants with a high and a low

need for autonomy, there was a significant main effect of the need for external control on perceived relevance, implying that participants with a high need for external control reported high positive levels of perceived relevance than respondents with a low need for external control. Table 5 provides full details of the results.

Table 5. Effect of choice, language use, and need for autonomy on perceived relevance for participants with a high and a low need for autonomy.

Variables	*Beta (B)*	SE *B*	*Beta (b)*	*t test (df)*	*P* value	95% CI
Low need for autonomy (n=242)						
Choice	−.05	0.05	−.05	−0.86 (7,235)	.39	−0.15 to .06
Language use	−.08	0.05	−.09	−1.40 (7,235)	.16	−0.18 to 0.03
Need for external control	.36	0.08	.28	4.58 (7,235)	<.001	0.21 to 0.52
Choice×language use	.01	0.05	.01	0.15 (7,235)	.89	−0.10 to 0.11
Choice×need for external control	.06	0.06	.06	1.00 (7,235)	.32	−0.06 to 0.18
Language×need for external control	−.12	0.06	−.12	−1.89 (7,235)	.06	−0.24 to 0.01
Choice×language×need for external control	−.09	0.06	−.09	−1.50 (7,235)	.13	−0.22 to 0.03
High need for autonomy (n=281)						
Choice	.15	0.06	.13	2.40 (7,274)	.02	1.74 to 2.59
Language use	.02	0.06	.02	0.34 (7,274)	.73	−0.10 to 0.14
Need for external control	.58	0.07	.44	8.20 (7,274)	<.001	0.44 to 0.72
Choice×language use	−.02	0.06	−.01	−0.25 (7,274)	.80	−0.14 to 0.10
Choice×need for external control	−.07	0.06	−.07	1.32 (7,274)	.19	−0.18 to 0.04
Language×need for external control	−.02	0.06	−.02	−0.38 (7,274)	.70	−0.13 to 0.09
Choice×language×need for external control	−.04	0.06	−.04	−0.68 (7,274)	.50	−0.15 to 0.07

Overall Evaluation

In terms of overall intervention rating, there was neither a significant interaction effect, rejecting hypotheses 3c and 4c, nor a significant main effect of language, rejecting hypothesis 1c. There was, however, a significant positive main effect of choice on the overall evaluation of the intervention (b=.12; P=.003). More specifically, the provision of choice was associated with a higher overall evaluation by the participants than no provision of choice, which confirms hypothesis 2c. Moreover, there were again main effects of the need for autonomy (b=.15; P<.001) and the need for external control (b=.32; P<.001), indicating that participants with a high need for autonomy and participants with a high need for external control evaluated the intervention significantly higher. Table 6 provides complete results of this analysis.

Table 6. Effect of choice, language use, and need for autonomy on overall evaluation of the intervention.

Variables	*Beta (B)*	SE *B*	*Beta (b)*	*t test (df)*	*P* value	95% CI
Choice	.14	0.05	.12	2.95 (11,513)	.003	0.05 to 0.24
Language use	−.04	0.05	−.03	−0.74 (11,513)	.46	−0.13 to 0.06
Need for autonomy	.27	0.08	.15	3.55 (11,513)	<.001	0.12 to 0.42
Need for external control	.49	0.06	.32	7.66 (11,513)	<.001	0.36 to 0.61
Choice×language use	−.00	0.05	−.00	−0.09 (11,513)	.93	−0.10 to 0.09
Choice×need for autonomy	.01	0.05	.01	−0.17 (11,513)	.87	−0.09 to 0.11
Choice×need for external control	−.07	0.05	−.06	−1.46 (11,513)	.15	−0.17 to 0.03
Language×need for autonomy	−.01	0.05	−.01	−0.24 (11,513)	.81	−0.11 to 0.09
Language×need for external control	.06	0.05	.05	1.22 (11,513)	.23	−0.04 to 0.16
Choice×language×need for autonomy	.06	0.05	.05	1.21 (11,513)	.23	−0.04 to 0.16
Choice×language×need for external control	−.07	0.05	−.06	−1.33 (11,513)	.19	−0.16 to 0.03

In a sensitivity analysis, we checked whether including a relative score of the need for autonomy (ie, score of need for autonomy–score of need for external control) as a potentially moderating variable, instead of 2 separate, unyoked variables for the need for autonomy and need for external control, yielded a change in results. The results were similar, yet 1 minor

difference was observed; the marginally significant effect of choice on more perceived autonomy-support turned nonsignificant (data not reported).

Discussion

Discussion of the Results

This study aimed to identify the most autonomy-supportive message frame within a Web-based computer-tailored intervention for increasing vegetable consumption. To this end, based on prior empirical research and theory, we investigated the effects of 2 strategies, that is, using autonomy-supportive language and offering choice, among Dutch adults on 3 outcomes (ie, perceived autonomy support, perceived relevance, and the overall evaluation of the intervention). Moreover, we examined whether individual differences in the need for autonomy and need for external control moderated these effects.

First of all, there appeared to be a main effect of choice on the overall evaluation of the intervention compared with no provision of choice, as well as positive effects of choice on perceived autonomy support that approached significance. This is in line with the expectations we had based on previous studies conducted in the face-to-face setting [9] and on the persuasive effects of choice more generally [29]. This can be considered preliminary evidence that the provision of choice could be an effective strategy to increase the effectiveness of Web-based computer-tailored health communication.

With regard to the potentially moderating role of need for autonomy, a significant interaction effect with choice was found for the dependent variable of perceived relevance: only for participants with a high need for autonomy was there a significant positive effect of choice. This is in line with our expectations, as we hypothesized that the positive effects of the provision of choice would be stronger for respondents with a high need for autonomy and strengthens the idea that message frame tailoring based on the need for autonomy might be a promising avenue to advance digital forms of health communication [6]. This idea is also empirically supported by findings from a recent study by the first author and colleagues, showing that customization—the ability to self-tailor the mediated environment [30,31]—in mobile health apps leads to higher intentions to engage in physical activity for those with a greater need for autonomy, but not for those with a smaller need for autonomy.

On the other hand, the language manipulation resulted in neither significant main effects nor significant interaction effects with the needs for autonomy and external control were found for any of the 3 dependent variables. Thus, the tone used in our health communication messages did not impact our 3 outcomes. It appeared not to matter whether the participants received the Web-based computer-tailored intervention using an autonomy-supportive or controlling communication style. This is, surprisingly, in line with the results from a recently published study into the effects of autonomy-supportive versus controlling message frames on individual's perceived autonomy support from and reactance toward such messages, studied in the context of a Web-based computer-tailored alcohol reduction intervention [15]. A potential explanation for this lack of effect—also mentioned by the authors from this recently published paper—might be derived from Politeness Theory [32]. In line with the operationalization we used in this study, Politeness Theory operationalizes autonomy-supportive language as not forceful language. In addition, however, this theory emphasizes the importance of perceived equality between the receiver and sender of the message. This implies that the receiver wants to be respected and seen as an equal by the sender of the message, or—in other words—their conversation partner, unless this conversation partner has a clear role of power over the receiver. If the conversation partner is perceived to be treating the receiver pedantically, although this is perceived to be outside of his or her power, the receiver might (un)consciously decide to resist the message and, consequently, not accept it—something also stressed by Psychological Reactance Theory [33]. As the source of the intervention was mentioned in the factsheet respondents received before entering the study and was intended to be perceived as expert (ie, 2 universities in combination with an innovative health consultancy agency), respondents in both conditions might have perceived the source of the intervention as having a clear role of power, resulting in the language manipulations no longer having any effect. To shed light on this, in future studies, it might be important to look at the moderating role of source characteristics and, for instance, investigate whether language and choice effects are different when the source of the message is (perceived as) an expert or a peer, as well as at resistance toward the message and/or source as a potential mediator of these effects.

Implications and Suggestions for Future Research

The provision of choice resulted in the Web-based computer-tailored intervention being more positively evaluated and perceived as more autonomy supportive compared with no provision of choice, although the effect of perceived autonomy support only approached significance and should be interpreted with caution. For health communication practice, this may imply that the provision of choice could be an effective and easy-to-implement strategy to increase the effectiveness of digital forms of health communication. Given its potentially large reach, this increased effectiveness may improve the impact of this low-cost health behavior change strategy on public health. There is, however, still room for future research in this area as, in this study, the provision of choice was operationalized by providing participants the possibility to indicate to what extent they wanted to make several preparatory and coping plans that were recommended based on previous research findings. This may be interpreted as a combination of both verbal choice (in this case, emphasizing to the respondents that they could choose which of the recommended plans they wanted to make) and physical choice (in this case, physically giving respondents the opportunity to indicate their desire to make a certain plan or not). As it has previously been suggested that different types of choice may lead to different effects [29], future research efforts might aim to disentangle the effects of these different types of choice and of different operationalizations of verbal and physical choice, respectively.

Second, the interaction effect found between choice and respondents' need for autonomy suggests that both health

behavior change theorists and health communication professionals may need to take each individual's communication style preferences (more) into account. Both SDT [7,12] and the I-Change Model [13] suggest the importance to tailor interventions to individual differences. SDT stresses the importance of need for autonomy in this regard, whereas the I-Change Model does not articulate which specific personal preferences need to be addressed as this is likely to depend on the topic, the target group, and the context. Consequently, the I-Change Model (ie, the theoretical framework used as a basis for the intervention studied) did not yet pay explicit attention to individual differences in the need for autonomy. The results from this study suggest that these sociocognitive models might actually benefit from doing so, suggesting an integration of theoretical concepts from different theories, a view also underlying the I-Change Model. For professionals, results suggest that when developing health communication interventions, people with a high need for autonomy—but not necessarily people with a low such need—might need to be provided with possibilities for choice throughout the intervention. In addition to the type of choice that was provided in this study (ie, the possibility to indicate to what extent participants wanted to make several preparatory and coping plans), other types of choice can be offered. For example, a recent study by the first author and colleagues shows that strategies such as customization and providing the respondent with explicit possibilities to choose could be used especially for people with a high need for autonomy. In contrast, system-driven tailoring strategies—tailoring based on an assessment of the individual respondent's characteristics, with no explicit choice options provided—might be better suited for respondents with a low need for autonomy. It should be noted, however, that the interaction effect described was only found for one of the dependent variables, that is, perceived relevance, and that for people with a low need for autonomy, no *negative* effect of choice was found. As a consequence, results should be interpreted with caution, and further research into the moderating role of the need for autonomy—as well as the need for external control—is warranted.

Strengths and Limitations

Some limitations also should be considered. First, we used perceived autonomy support, perceived relevance, and the overall evaluation of the intervention as outcome measures, assessed directly postintervention. Although we can assume that these measures would facilitate the internalization of motivation and ultimately predict health behavior change and its maintenance [7], future research efforts might want to consider measuring respondents' motivation, intention to change, and/or actual behavior (change) using a longitudinal research design with a longer follow-up period. Second, we tested the effects of language and choice in the context of a Web-based computer-tailored intervention only. Although this was based on evidence from several previous studies [1] and we wanted to test the add-on effect of variations in message frame, future research might aim at testing whether the theoretical hypotheses also hold—may in fact, be stronger—in a nontailored context. Third, as participants in the choice conditions received feedback that was tailored based on the preparatory and coping plans they chose (not) to make and participants who were in the no-choice conditions received 1 (nontailored) advice statement for both preparatory planning and coping planning, it might be difficult to disentangle the effect of providing choice from the effect of an additional amount of content-tailored feedback. Future research could operationalize choice independently from the content tailoring, for example, by letting respondents choose whether and for which out of a predetermined set of potentially difficult situations they want to make a coping plan and comparing this with respondents instructed to make plans for all these situations—not providing them with any subsequent (content-tailored) feedback, or by letting respondents choose for a date at which they agree to start their lifestyle change, for example, to start eating more vegetables or to stop smoking versus providing respondents with a randomly generated start date. Finally, only respondents interested in eating—or continuing to eat—250 g of vegetables per day were included in the study. This inclusion criterion was deliberately chosen. As a positive motivation to change has been acknowledged by theory (eg, studies by de Vries et al [13] and Ajzen [34]) and evidence (eg, studies by Vangeli et al [35] and Smit et al [36]) to be a necessary prerequisite for actual behavior change, future research efforts might additionally explore the effects of language and especially choice among less motivated individuals.

Conclusions

This study suggests that provision of choice rather than the use of autonomy-supportive language can be an easy-to-implement strategy to increase the effectiveness of digital health communication, especially for people with a high need for autonomy.

Acknowledgments

This study was supported by the Innovational Research Incentives Scheme Veni from NWO-MaGW (Netherlands Organization for Scientific Research—Division for the Social Sciences) accredited to ESS (project number 451-15-028).

Conflicts of Interest

None declared.

Multimedia Appendix 1
Consolidated Standards of Reporting Trials flow diagram.

[PDF File (Adobe PDF File), 66 KB - jmir_v21i10e14074_app1.pdf]

Multimedia Appendix 2
An example of a feedback message in autonomy-supportive and controlling language combined with and without choice.
[PDF File (Adobe PDF File), 75 KB - jmir_v21i10e14074_app2.pdf]

Multimedia Appendix 3
CONSORT-EHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 75 KB - jmir_v21i10e14074_app3.pdf]

References

1. Lustria ML, Noar SM, Cortese J, van Stee SK, Glueckauf RL, Lee J. A meta-analysis of web-delivered tailored health behavior change interventions. J Health Commun 2013;18(9):1039-1069. [doi: 10.1080/10810730.2013.768727] [Medline: 23750972]
2. Schulz DN, Smit ES, Stanczyk NE, Kremers SP, de Vries H, Evers SM. Economic evaluation of a web-based tailored lifestyle intervention for adults: findings regarding cost-effectiveness and cost-utility from a randomized controlled trial. J Med Internet Res 2014;16(3):e91 [FREE Full text] [doi: 10.2196/jmir.3159] [Medline: 24650860]
3. Cheung KL, Wijnen B, de Vries H. A review of the theoretical basis, effects, and cost effectiveness of online smoking cessation interventions in the Netherlands: a mixed-methods approach. J Med Internet Res 2017 Jun 23;19(6):e230 [FREE Full text] [doi: 10.2196/jmir.7209] [Medline: 28645889]
4. Glasgow RE, Vogt TM, Boles SM. Evaluating the public health impact of health promotion interventions: the RE-AIM framework. Am J Public Health 1999 Sep;89(9):1322-1327. [doi: 10.2105/AJPH.89.9.1322] [Medline: 10474547]
5. Rimer BK, Kreuter MW. Advancing tailored health communication: a persuasion and message effects perspective. J Communication 2006 Aug;56(s1):S184-S201. [doi: 10.1111/j.1460-2466.2006.00289.x]
6. Smit ES, Linn AJ, van Weert JC. Taking online computer-tailoring forward: the potential of tailoring the message frame and delivery mode of online health behaviour change interventions. Eur Health Psychol 2015;17(1):25-31 [FREE Full text]
7. Ng JY, Ntoumanis N, Thøgersen-Ntoumani C, Deci EL, Ryan RM, Duda JL, et al. Self-determination theory applied to health contexts: a meta-analysis. Perspect Psychol Sci 2012 Jul;7(4):325-340. [doi: 10.1177/1745691612447309] [Medline: 26168470]
8. Patrick H, Canevello A. Methodological overview of a self-determination theory-based computerized intervention to promote leisure-time physical activity. Psychol Sport Exerc 2011 Jan;12(1):13-19 [FREE Full text] [doi: 10.1016/j.psychsport.2010.04.011] [Medline: 21103069]
9. Deci EL, Eghrari H, Patrick BC, Leone DR. Facilitating internalization: the self-determination theory perspective. J Pers 1994 Mar;62(1):119-142. [doi: 10.1111/j.1467-6494.1994.tb00797.x] [Medline: 8169757]
10. Williams GC, Cox EM, Kouides R, Deci EL. Presenting the facts about smoking to adolescents: effects of an autonomy-supportive style. Arch Pediatr Adolesc Med 1999 Sep;153(9):959-964. [doi: 10.1001/archpedi.153.9.959] [Medline: 10482213]
11. Williams GC, Niemiec CP, Elliot AJ, LaGuardia JG, Gorin AA, Rigby CS. Virtual Look AHEAD program: initial support for a partly virtualized intensive lifestyle intervention in type 2 diabetes. Diabetes Care 2014 Aug;37(8):e169-e170 [FREE Full text] [doi: 10.2337/dc14-0831] [Medline: 25061144]
12. Deci EL, Ryan RM. The general causality orientations scale: self-determination in personality. J Res Pers 1985 Jun;19(2):109-134. [doi: 10.1016/0092-6566(85)90023-6]
13. de Vries H, Mudde A, Leijs I, Charlton A, Vartiainen E, Buijs G, et al. The European smoking prevention framework approach (EFSA): an example of integral prevention. Health Educ Res 2003 Oct;18(5):611-626 [FREE Full text] [doi: 10.1093/her/cyg031] [Medline: 14572020]
14. Gillison FB, Rouse P, Standage M, Sebire SJ, Ryan RM. A meta-analysis of techniques to promote motivation for health behaviour change from a self-determination theory perspective. Health Psychol Rev 2019 Mar;13(1):110-130. [doi: 10.1080/17437199.2018.1534071] [Medline: 30295176]
15. Altendorf MB, van Weert JC, Hoving C, Smit ES. Should or could? Testing the use of autonomy-supportive language and the provision of choice in online computer-tailored alcohol reduction communication. Digit Health 2019;5:2055207619832767 [FREE Full text] [doi: 10.1177/2055207619832767] [Medline: 30834136]
16. de Vries H, Brug J. Computer-tailored interventions motivating people to adopt health promoting behaviours: introduction to a new approach. Patient Educ Couns 1999 Feb;36(2):99-105. [Medline: 10223015]
17. Strecher VJ, McClure JB, Alexander GL, Chakraborty B, Nair VN, Konkel JM, et al. The role of engagement in a tailored web-based smoking cessation program: randomized controlled trial. J Med Internet Res 2008;10(5):e36 [FREE Full text] [doi: 10.2196/jmir.1002] [Medline: 18984557]
18. Dijkstra A. Working mechanisms of computer-tailored health education: evidence from smoking cessation. Health Educ Res 2005 Oct;20(5):527-539 [FREE Full text] [doi: 10.1093/her/cyh014] [Medline: 15701665]

19. Levinson W, Kao A, Kuby A, Thisted RA. Not all patients want to participate in decision making. A national study of public preferences. J Gen Intern Med 2005 Jun;20(6):531-535 [FREE Full text] [doi: 10.1111/j.1525-1497.2005.04101.x] [Medline: 15987329]
20. Resnicow K, Zhou Y, Hawley S, Jimbo M, Ruffin MT, Davis RE, et al. Communication preference moderates the effect of a tailored intervention to increase colorectal cancer screening among African Americans. Patient Educ Couns 2014 Dec;97(3):370-375 [FREE Full text] [doi: 10.1016/j.pec.2014.08.013] [Medline: 25224317]
21. Kreuter MW, Farrell DW, Olevitch LR, Brennan LK. Tailoring Health Messages: Customizing Communication with Computer Technology. New York: Lawrence Erlbaum Associates; 1999.
22. Resnicow K, Davis RE, Zhang G, Konkel J, Strecher VJ, Shaikh AR, et al. Tailoring a fruit and vegetable intervention on novel motivational constructs: results of a randomized study. Ann Behav Med 2008 Apr;35(2):159-169. [doi: 10.1007/s12160-008-9028-9] [Medline: 18401673]
23. Schulz DN, Kremers SP, Vandelanotte C, van Adrichem MJ, Schneider F, Candel MJ, et al. Effects of a web-based tailored multiple-lifestyle intervention for adults: a two-year randomized controlled trial comparing sequential and simultaneous delivery modes. J Med Internet Res 2014;16(1):e26 [FREE Full text] [doi: 10.2196/jmir.3094] [Medline: 24472854]
24. PanelClix - Het grootste Panel van Nederland. URL: http://www.panelclix.nl [accessed 2016-06-15] [WebCite Cache ID 6iHJ2h3dX]
25. Brink L, Postma-Smeets A, Stafleu A, Wolvers D. The Netherlands Nutrition Centre Foundation. 2016. Richtlijnen Schijf van Vijf. URL: https://www.voedingscentrum.nl/Assets/Uploads/voedingscentrum/Documents/Professionals/Schijf%20van%20Vijf/Voedingscentrum%20Richtlijnen%20Schijf%20van%20Vijf%202016%204.pdf [accessed 2019-09-05]
26. Williams GG, Gagné M, Ryan RM, Deci EL. Facilitating autonomous motivation for smoking cessation. Health Psychol 2002 Jan;21(1):40-50. [doi: 10.1037/0278-6133.21.1.40] [Medline: 11846344]
27. National Institute for Public Health and the Environment. 2005. [Indicators of the Public Health Monitor] URL: https://www.monitorgezondheid.nl/volksindicatoren.aspx [accessed 2017-07-12] [WebCite Cache ID 6rtlcEUDn]
28. Smit ES, Dima AL, Immerzeel SA, van den Putte B, Williams GC. The virtual care climate questionnaire: development and validation of a questionnaire measuring perceived support for autonomy in a virtual care setting. J Med Internet Res 2017 May 8;19(5):e155 [FREE Full text] [doi: 10.2196/jmir.6714] [Medline: 28483743]
29. Patall EA, Cooper H, Robinson JC. The effects of choice on intrinsic motivation and related outcomes: a meta-analysis of research findings. Psychol Bull 2008 Mar;134(2):270-300. [doi: 10.1037/0033-2909.134.2.270] [Medline: 18298272]
30. Kalyanaraman S, Sundar SS. The psychological appeal of personalized content in web portals: does customization affect attitudes and behavior? J Commun 2006 Mar;56(1):110-132. [doi: 10.1111/j.1460-2466.2006.00006.x]
31. Sundar SS, Bellur S, Jia H. Motivational Technologies: A Theoretical Framework for Designing Preventive Health Applications. In: Proceedings of the International Conference on Persuasive Technology. 2012 Presented at: PERSUASIVE'12; June 6-8, 2012; Linköping, Sweden p. 112-122. [doi: 10.1007/978-3-642-31037-9_10]
32. Jenkins M, Dragojevic M. Explaining the process of resistance to persuasion. Commun Res 2011;40(4):559-590 [FREE Full text] [doi: 10.1177/0093650211420136]
33. Brehm JW. A Theory of Psychological Reactance. Oxford, England: Academic Press; 1966.
34. Ajzen I. From intentions to actions: a theory of planned behavior. In: Kuhl J, Beckmann J, editors. Action Control: From Cognition to Behavior. Heidelberg, Germany: Springer; 1985:11-39.
35. Vangeli E, Stapleton J, Smit ES, Borland R, West R. Predictors of attempts to stop smoking and their success in adult general population samples: a systematic review. Addiction 2011 Dec;106(12):2110-2121. [doi: 10.1111/j.1360-0443.2011.03565.x] [Medline: 21752135]
36. Smit ES, Hoving C, Schelleman-Offermans K, West R, de Vries H. Predictors of successful and unsuccessful quit attempts among smokers motivated to quit. Addict Behav 2014 Sep;39(9):1318-1324. [doi: 10.1016/j.addbeh.2014.04.017] [Medline: 24837754]

Abbreviations

BMI: body mass index
HCOS: Health Causality Orientation Scale
SDT: Self-Determination Theory

Edited by G Eysenbach; submitted 20.03.19; peer-reviewed by KL Cheung, S Lippke; comments to author 14.06.19; revised version received 20.08.19; accepted 20.08.19; published 30.10.19.

Please cite as:
Smit ES, Zeidler C, Resnicow K, de Vries H
Identifying the Most Autonomy-Supportive Message Frame in Digital Health Communication: A 2x2 Between-Subjects Experiment
J Med Internet Res 2019;21(10):e14074
URL: http://www.jmir.org/2019/10/e14074/
doi:10.2196/14074
PMID:31670693

Original Paper

Home Virtual Visits for Outpatient Follow-Up Stroke Care: Cross-Sectional Study

Ramana Appireddy[1], MBBS, DM, MSc; Sana Khan[1], HBsc, MSc; Chad Leaver[2], BA, MSc, MBA; Cally Martin[3], BSc, PT, MSc; Albert Jin[1], PhD, MD, FRCP(C); Bryce A Durafourt[1], MDCM, MSc; Stephen L Archer[4], MD, FRCP(C)

[1]Division of Neurology, Department of Medicine, Kingston Health Sciences Centre, Kingston, ON, Canada

[2]Canada Health Infoway, Toronto, ON, Canada

[3]Stroke Network of Southeastern Ontario, Kingston Health Sciences Centre, Kingston, ON, Canada

[4]Department of Medicine, Kingston Health Sciences Centre, Kingston, ON, Canada

Corresponding Author:
Ramana Appireddy, MBBS, DM, MSc
Division of Neurology
Department of Medicine
Kingston Health Sciences Centre
Kingston General Hospital, Connell 7
76 Stuart Street
Kingston, ON, K7L 2V7
Canada
Phone: 1 613 549 6666 ext 6320
Email: mrra@queensu.ca

Abstract

Background: Timely, in-person access to health care is a challenge for people living with conditions such as stroke that result in frailty, loss of independence, restrictions in driving and mobility, and physical and cognitive decline. In Southeastern Ontario, access is further complicated by rurality and the long travel distances to visit physician clinics. There is a need to make health care more accessible and convenient. Home virtual visits (electronic visits, eVisits) can conveniently connect physicians to patients. Physicians use a secure personal videoconferencing tool to connect to patients in their homes. Patients use their device of choice (smartphone, tablet, laptop, or desktop) for the visit.

Objective: This study aimed to assess the feasibility and logistics of implementing eVisits in a stroke prevention clinic for seniors.

Methods: A 6-month eVisit pilot study was initiated in the Kingston Health Sciences Centre stroke prevention clinic in August 2018. eVisits were used only for follow-up patient encounters. An integrated evaluation was used to test the impact of the program on clinic workflow and patient satisfaction. Patient satisfaction was evaluated by telephone interviews, using a brief questionnaire. Access and patient satisfaction metrics were compared with concurrent standard of care (patients' prior personal experience with in-person visits). Values are presented as median (interquartile range).

Results: There were 75 subjects in the pilot. The patients were aged 65 (56-73.5) years, and 39% (29/75) resided in rural areas. There was a shorter wait for an appointment by eVisit versus in-person (mean 59.98 [SD 48.36] days vs mean 78.36 [SD 50.54] days; $P<.001$). The eVisit was also shorter, taking on an average of only 10 min to deliver follow-up care with a high degree of patient satisfaction versus 90 (60-112) min for in-person care. The total time saved by patients per eVisit was 80 (50-102) min, 44 (21-69) min of which was travel time. Travel distance avoided by the patients was 30.1 km (11.2-82.2). The estimated total out-of-pocket cost savings for patients per eVisit was Can $52.83 (31.26-94.53). The estimated savings (opportunity cost for in-person outpatient care) for our eVisit pilot project was Can $23,832-$28,584. The patient satisfaction with eVisits was very good compared with their prior personal experience with in-person outpatient care.

Conclusions: The eVisit program was well received by patients, deemed to be safe by physicians, and avoided unnecessary patient travel and expense. It also has the potential to reduce health care costs. We plan to scale the project within the department and the institution.

(J Med Internet Res 2019;21(10):e13734) doi:10.2196/13734

KEYWORDS

telemedicine; eHealth; eVisit; mobile health; health services accessibility

Introduction

Barriers to Care

Canadians face many barriers while accessing outpatient health care services, including accessibility, availability, acceptability, and personal choice [1,2]. In a recent survey of the western Canadian provinces, 10% of adults with chronic health conditions reported having barriers to accessing outpatient primary care [2]. The overall health care experience was found to be poor in patients with chronic health conditions even when they reported good access because of a perceived failure of the system to meet their needs [3]. Similar barriers to outpatient specialist care exist and particularly impact seniors (aged 65 years and older) because of their higher prevalence of chronic health conditions and frailty. Seniors are also more likely to have reduced functional capacity, lower socioeconomic status, reduced independence, cognitive decline, and driving restrictions [4-6]. Patient-centered care has become a critical component of health policy worldwide and is best summarized by the Picker's principles of patient-centered care [7]. To optimize health system performance, the Institute for Healthcare Improvement developed the triple aim framework, the goals of which are to improve the patient care experience (including quality and satisfaction), improve the health of populations, and reduce the per capita cost of health care [8]. Evidence suggests that health system transformation needs to be reformed from a patient-centric perspective to meet the health care needs of seniors [4,9,10].

Telemedicine and Virtual Visits

Traditional practice utilizes in-person interaction to establish the patient-physician relationship and to complete a comprehensive clinical evaluation, including history and physical examination. However, follow-up care, including symptom management, diagnostics, and therapeutic decision making require less in-person interaction and may be achieved by virtual visits. Virtual visits, also known as eVisits, are a secure, 2-way digital communication between health providers and their patients. eVisits may include emails, short message service text messaging, and videoconferencing [11]. A recent study from British Columbia suggests that virtual visits in primary care are associated with a high degree of patient satisfaction and positive system outcomes [12]. There is patient demand for electronic health services. In a 2018 national survey, only 6% of Canadians over the age of 16 years said that they could currently visit their health care provider online by video, whereas among those who could not, 47% desired such access [13]. Health care professionals also perceive that offering health care that is convenient to the patient is an essential aspect of good clinical medicine [14]. There is also a growing call for virtualization of health care by health care professionals, policy makers, and industry leaders [14-16].

eVisits and other telemedicine modalities across Ontario, Canada, are facilitated by technologies provided by the Ontario Telemedicine Network (OTN), a not-for-profit organization funded by the Ontario Ministry of Health and Long-Term Care (MOHLTC). For the purposes of this publication, "eVisit" refers to personal, secure videoconferencing between the health care provider and the patient. Unlike conventional telehealth modalities, an eVisit does not need new infrastructure, such as dedicated videoconferencing equipment, peripheral devices, or a telemedicine facility, and the patient remains in their home. In an eVisit, the physician and patient interface using electronic equipment that is widely available, such as smartphones and tablets. The traditional telemedicine model with the patient at a remote site (satellite site) reduces the patient travel burden but is still costly to the health care system as significant infrastructure is used at both ends. eVisits have been shown to be feasible, acceptable, and yield similar clinical outcomes compared with in-person patient cohorts in an interdisciplinary obesity treatment program for adolescents in Ontario, Canada [17].

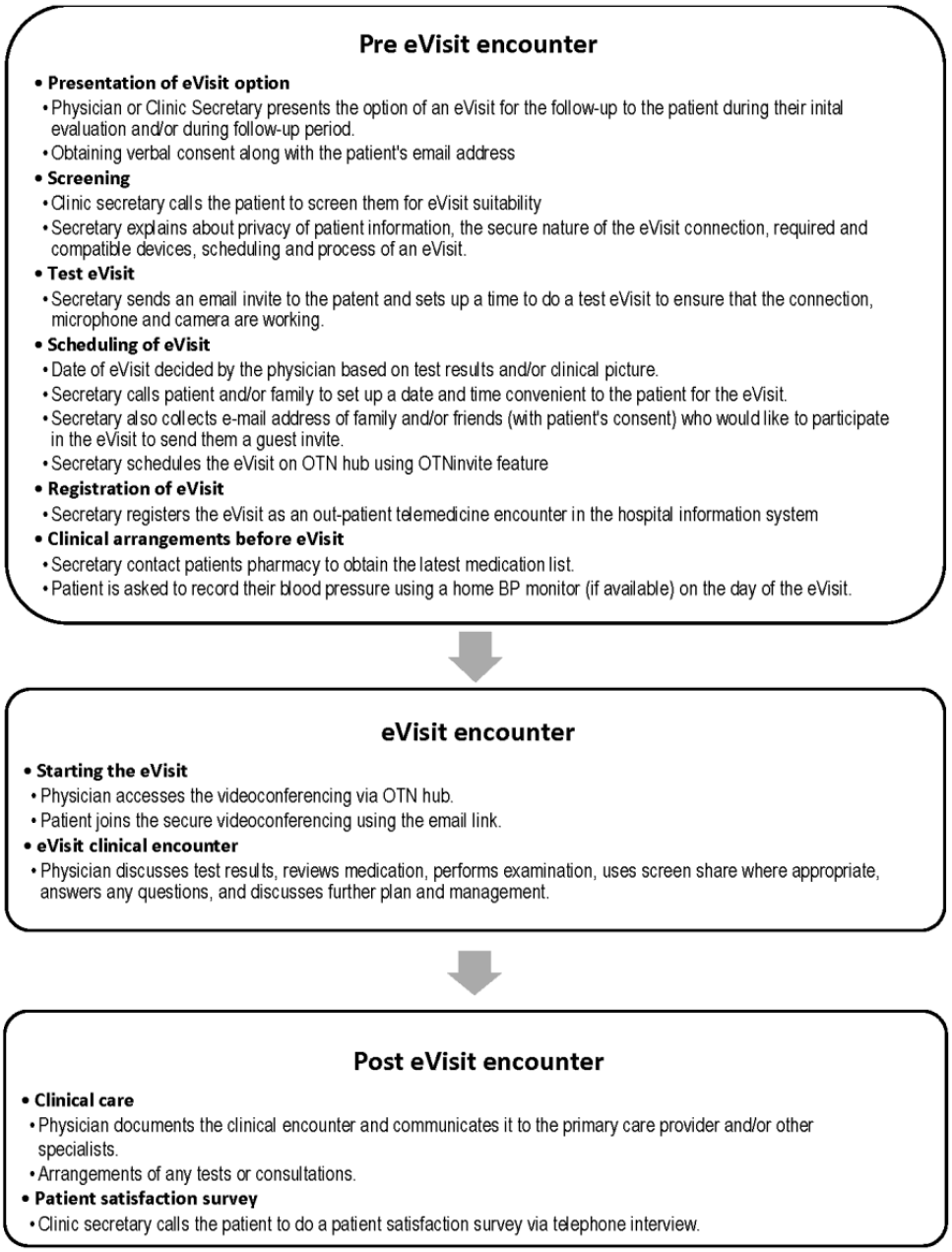
Whether eVisits would also be beneficial to seniors, a group traditionally viewed as being less technologically adept, was tested in a 6-month pilot project in a high-volume stroke prevention clinic in Ontario at the Kingston Health Sciences Centre (KHSC). The results of this pilot program indicated that the eVisit was well received by patients and has the potential to provide cost savings to both patients and the health care system.

Methods

Study Setting

The eVisit pilot study was initiated at KHSC in August 2018 for a 6-month period with the objective to assess the feasibility and logistics of implementing eVisits in an adult specialty disease clinic catering predominantly to seniors. An integrated evaluation was designed to test the impact of the pilot program both on clinic workflow and patient satisfaction using a telephone survey. The workflow of the eVisit intervention in the stroke prevention clinic is presented in Figure 1. The selection criteria used for an eVisit are presented in Textboxes 1 and 2. The eVisit was done through a secure Web platform hosted by OTN. The physician used a desktop computer equipped with a Web camera and a microphone for all the eVisit encounters. The patients used the device of their preference. In addition, the eVisit platforms offer other features including a guest invitation option, wherein up to 6 more participants (family, friends, or other health care team members) can join in the videoconference.

Figure 1. Electronic visit workflow. BP: blood pressure; OTN: Ontario Telemedicine Network.



Textbox 1. Inclusion criteria for patients suitable for an independent electronic visit: patient characteristics.

- No cognitive issues.
- No loss of communication abilities.
- No physical deficits and loss of functional abilities.
- No sensory or perceptual deficits.
- No visual field deficits with functional implications.

Textbox 2. Inclusion criteria for patients suitable for an independent electronic visit: electronic visit technical eligibility.

- Patient/substitute decision maker (SDM) willing to do an electronic visit (eVisit) for follow-up care.
- Patient/SDM has internet-enabled device (smartphone, tablet, or computer).
- Patient/SDM has access to an internet connection.
- Patient/SDM has a secure place to perform an eVisit.

The study subjects were selected from patients routinely seen in the stroke prevention clinic using prespecified criteria (Textboxes 1 and 2). This was used only for follow-up appointments, and the characteristics of the patients, clinical or eVisit characteristics, patient satisfaction survey, and the impact on the clinic wait times are described. The patient satisfaction survey asked the patients to report their experience with eVisits compared with their personal experience with prior in-person health care; thus, the subjects acted as their own controls. Follow-up wait times for patients seen by an eVisit were compared with a control arm consisting of patients seen for in-person follow-up in the stroke prevention clinic over the same time period. The follow-up wait time is measured as the time from initial evaluation to follow-up appointment.

eVisits were used only for the follow-up visits. Patients had to fulfill the *patient characteristics* and *eVisit technical eligibility* criteria as shown in Textboxes 1 and 2. Individualized decisions were made for patients who were willing but did *not* meet *patient characteristics* eligibility *and* had a family member or substitute decision maker that fulfilled the eVisit technical eligibility.

The definition for *senior* status used for this study is *aged 65 and over* as adapted from Statistics Canada [6]. The rural residence is defined based on the second character of the 6-digit postal code of the patient's home address [18]. Savings on patient time (total time, travel time, and in-person visit time) and travel distance avoided were estimated using these definitions (Multimedia Appendix 1). The in-person visit time was conservatively estimated at 30 min in addition to the travel time.

Methodology for Economic Analysis

We also performed a preliminary economic analysis to estimate hypothetical out-of-pocket (OOP) patient cost savings of eVisits and opportunity costs of in-person outpatient care. Opportunity cost are defined as benefits foregone by the particular use of resources, resources which could be otherwise allocated for other health care priorities [19]. These are hypothetical estimates based on some assumptions mentioned below and not based on actual economic data (income, employment status, outpatient costing, etc) from the patients in the pilot or the hospital.

Out-of-Pocket Cost Estimate

The OOP expenses were not captured using a specially designed survey; however, approximate reasonable OOP expenses were estimated using the cost of travel, parking, potential loss of pay, and total cost (Multimedia Appendix 1). As there are no current data on the impact of virtual care on OOP expenses for outpatient care, we attempted to estimate this in our study. The potential loss of pay for an adult obtaining in-person care is estimated using a hypothetical estimate of what an adult Canadian older than 25 years working full-time would lose on average if they had visited a doctor for an in-person visit, assuming the travel distance and time spent per in-person visit are similar to the study cohort [20]. The loss of pay was estimated using the average full-time hourly wage of Can $28.98 per hour for Canadian adults aged 25 years and older based on 2018 Statistics Canada data [20]. Additional OOP expenses such as food expenses, childcare costs, loss of pay for a family member, or other caregiver costs were not considered for our analysis of OOP estimates.

Outpatient Hospital Cost and Opportunity Cost of In-Person Outpatient Hospital Care

Our institution does not collect or report outpatient costing to Canadian Institute for Health Information (CIHI), so we used the available provincial outpatient costing data for reference [21-23]. The outpatient or ambulatory care costs are reported only by a few hospitals across the country and include the direct costs (nursing, diagnostic tests, operating, and recovery room), functional center indirect costs (meals, facilities management, and plant operation), and costs for patient-specific drugs and supplies. Using the Comprehensive Ambulatory Classification System (CACS) developed by CIHI for 2 codes (E751: General Signs, Symptoms, Examinations and Investigations; E752: Other Medical and Follow-up Care), the outpatient hospital costs in Alberta (interactive health data application) and Ontario (Ontario Case Costing Initiative Costing Analysis Tool) for 2016/17 were used as reference for estimating the outpatient hospital costs for our eVisit cohort [21-23].

Statistical Analysis

Data were entered into Epidata software (The Epidata Association, Denmark) and were analyzed using STATA v15.0 (StataCorp LLC, USA). The data was analyzed using summary statistics and Wilcoxon Signed Rank Sum test for paired data. Values are stated as the mean and interquartile range (IQR). A $P<.05$ is considered statistically significant.

Ethics Approval

Ethics approval for the pilot study was obtained, permitting for chart review and data collection (Queen's University Ethics ROMEO # 6025439).

Results

Baseline Data

There were a total of 75 eVisits from August 2018 to January 2019. The details on the overall clinic volumes and the number of eVisits are provided in the Multimedia Appendix 2. A formal screening log was not maintained; however, some of the factors that influenced the patient uptake include lack of interest, lack of technology, as well as physician engagement in offering eVisits. Promotion of the eVisit by the physician resulted in higher uptake compared with engagement by the clinic secretary. During the pilot project, 40.2% (76/189) of the follow-up visits were through eVisit.

The mean (SD) and median (IQR) age of the patients was 63.7(14.3) and 65 years (56-73.5), respectively. Of the study patients, 67% (50/75) were male, 51% (38/75) were under age 65, 32% (24/75) were aged 65-75, and 17% (13/75) were over age 75. Mobile internet devices, including tablets (68%; 51/75) and smartphones (24%; 18/75), were most widely used for the eVisits, likely because of ease of use and setup. Laptops were used for 7% (5/75) and desktops were used for 1% (1/75) of eVisits. The mean (SD) and median (IQR) time spent by the physician and the patient for an eVisit encounter was 9.81 (4.46)

and 10 (9-12) min, respectively. The proportion of rural residents who performed eVisits was 39% (29/75). A single family member accompanied the patient in 60% (45/75) of the eVisit encounters. A total of 11% (8/75) of the patients were at their place of work for the eVisit, the eVisit was done in a secure location selected by the patient, and none of them needed to take time off work.

The Wilcoxon signed-rank test showed significant reduction (P<.001) in the mean (SD) wait times for follow-up for in-person (mean 78.36 days, SD 50.54 days) compared with eVisit follow-up (mean 59.98 days, SD 48.36 days).

The savings on travel avoided, time savings, and direct patient OOP expenses are presented in Multimedia Appendix 3. The median value for total time saved and total travel distance avoided are 80 (50-102) min and 30.1 (11.2-82.2) km, respectively.

The various components of patient care during the eVisit included, when relevant, a review of imaging tests (33%; 24/72), cardiac tests (43%; 31/72), lab tests (26%; 19/72), consults from other specialists (28%; 20/72), medication reconciliation (93%; 67/72), and potential new tests or interventions (50%; 36/72). Screen sharing was used for 28% (20/72) of eVisits. The diagnosis at the time of eVisit included stroke (49%; 35/72), transient ischemic attack (33%; 24/72), migraine (3%; 2/72), epilepsy (8%; 6/72), or other (7%; 5/72).

A telephone patient experience survey was also completed by patients that had an eVisit with a good survey response (46%; 33/72). The degree of patient satisfaction captured using the survey questionnaire was very high (Table 1). Almost all of the respondents agreed that the eVisits saved them time (100%;33/33), money (97%;32/33), and avoided traveling to the doctor's office (100%;33/33). All the patients who responded reported having a better experience via eVisit compared with an in-person visit and felt that their health issue was appropriately addressed during the eVisit. All the patients that had an eVisit were very willing to use the eVisit for further follow-up encounters. More than 90% (31/33) of the patients reported that they would strongly recommend an eVisit option or process to their friends and family. This was the first virtual health care encounter for all the patients that were involved in the eVisit pilot program. All were pleased with the convenience it offered and shared their experiences on a voluntary basis and were not enrolled in a formal qualitative study (Textbox 3).

Table 1. Patient experience from electronic visits (N=33).

Question	n (%)
Did eVisit[a] save you time?	
Yes	33 (100)
Did eVisit save you money?	
Yes	32 (97)
Did eVisit allow to avoid traveling to your doctor or specialist?	
Yes	33 (100)
Do you think if your health issue was addressed appropriately during the eVisit?	
Yes	33 (100)
Did you feel that the security and privacy of your health care information were protected during the eVisit?	
Yes	33 (100)
How is the experience of care from using the eVisit compared with an in-person encounter?	
Better	12 (36)
Same	19 (58)
Not sure	2 (6)
Worse	0 (0)
Would you use eVisit again?	
Definitely	31 (94)
Probably	2 (6)
Neutral, probably not, definitely not, not sure	0 (0)
How likely are you to recommend the eVisit to a friend on a scale of 0-10?	
0-7	2 (6)
8-100	31 (94)

[a]eVisit: electronic visit.

Economic Analysis

The estimates for OOP costs saved are presented in Table 2. The median estimate for total OOP patient cost savings by using the eVisit instead of the in-person visit was Can $52.83 (31.26-94.53).

The outpatient hospital-based health care costs in Ontario and Alberta for codes E751 and E752 based on CIHI CACS for 2016/17 are provided in the Multimedia Appendix 4 [21-23]. Using the Ontario data as the reference, the estimated cost of outpatient hospital care for our eVisit cohort was between Can $23,832 to 28,584.

Textbox 3. Excerpts of written feedback received from patients and family about the electronic visit experience.

"It is really important for a patient to have a proper conversation with their doctor, to ask questions, and to get answers which put my mind at rest. I felt as if I was in your office talking to you face to face. Not having to arrange transportation to get to your office was a real help. Now that I am not able to drive, mobility within the community is a real issue. I hope that this will be something you can offer to patients regularly." [Female patient, aged 90 years]

"The e-visit saved time and a lot of stress that is involved in taking an elderly patient out especially in bad weather." [Family of a male patient aged 88 years]

"It was nice not to have had to drive to the hospital, pay for parking, and make the physical effort of getting to the appointment destination." [Male patient, aged 88 years]

Table 2. Estimated out-of-pocket cost savings to patients in Canada.

Out-of-pocket expenses category	Total (Can $)	Can $, mean (SD)	Can $, median (interquartile range)
Patient self-reported cost for in-person visit (n=24)	417	13.4 (14.5)	10 (5-15)
Estimated travel cost for patients	2384.46	33.13 (36.92)	16.55 (6.16-45.21)
Estimated total out-of-pocket savings	5393.97	74.92 (57.99)	52.83 (31.26-94.53)

Discussion

Electronic Visit Implementation

We demonstrated that eVisits could be successfully implemented for secondary prevention of stroke in an adult neurology clinic catering predominantly to seniors. eVisits are time-efficient for physicians and patients, taking a median time of 10 min while avoiding the logistical challenges of an in-person encounter and reducing OOP expenses. Patient satisfaction is very high with the eVisits. A significant proportion (33%;11/33) of our cohort reported the experience to be better than an in-person encounter. During the eVisit, it is possible to perform most of the conventional components of clinical care that happen during a routine follow-up clinic visit for this patient population. The proportion of cancellations and no-shows is minimal, highlighting the impact of the eVisits on the overall efficiency of this model of ambulatory care. There is a significant reduction in the wait times for the patients via eVisit compared with in-person follow-up, which is likely because of the lack of the need for traditional health care infrastructure. The direct translatable savings to the patients with regard to OOP expenses for travel avoided and time saved are substantial.

Conventional Telemedicine Versus Electronic Visits

There is extensive literature describing the positive impact of conventional telemedicine modalities on access to health care globally [24-28]. Conventional telemedicine modalities such as remote videoconferencing between a host site and peripheral telemedicine satellite site have reduced the need for patient travel, reduced wait times, and improved coverage. However, these conventional models still require the patient to go to the satellite telemedicine site and are expensive, requiring significant specialized infrastructure and personnel at the satellite site. eVisits refine and simplify telehealth (for appropriate applications) by keeping the patient at home, reducing the need for any travel, and eliminating the need for specialized health care infrastructure.

Use of Virtual Visits

The use of eVisits has grown in the last few years across North America [29]. Although this model is embraced by patients, physicians, insurance providers, and policy makers, the virtual care model is used for a few conditions in primary care [14,15,30,31] but is not broadly utilized by Canadian family physicians [32]. Broadly, there are 2 models of eVisit (virtual visit) currently available in Canada and the United States—a pay per use model and an insured model. In a pay per use model, also referred to as Direct to Consumer, the consumer (patient and/or family) pays a fee to access a physician for a health care consultation through a virtual visit [33]. In the insured model, virtual visits are covered by private or public health insurance. They are predominantly used in primary care with the majority of the use restricted to routine or common primary health conditions or situations including the common cold, skin rash, and prescription renewals [12]. Published reports of eVisits (home virtual visits) is limited to certain specific diseases or conditions such as acute respiratory illnesses or urinary tract symptoms [34-38]. Overall, specialist use of virtual visits is higher (9%) compared with primary care (4%) [32]. The use of virtual visits for specialist care is limited in Canada, with the predominant use in psychiatry (personal communication with OTN). Thus, our experience is one of the first reports on the use of eVisits by specialists in Canada.

Electronic Visits Address Triple Aim and Patient-Centered Care

Patient Experience of Care

eVisits were associated with high patient satisfaction when employed in primary care settings [12]. The KHSC pilot demonstrated that the eVisits offer a high degree of patient satisfaction in specialty care among seniors and align with Picker's principles of patient-centered care [7]. In our study, seniors (aged 65 and older) constituted 49% (37/75) of the eVisit clientage, suggesting an increased acceptance of virtual care in this group. The results from the patient experience survey suggest that the eVisit process positively addressed patients' perceptions of accessibility and acceptability. The impression that an eVisit is better than an in-person encounter reported by 33% (11/33) of stroke patients in our pilot program is likely because of ease of facilitating an in-home follow-up consultation, avoiding the time, effort, or stress of arranging transportation, travel, parking, waiting in the clinic, loss of pay, need for caregiver assistance, and associated time savings. Another major advantage of the eVisits is avoiding the significant influence of weather and road conditions on commuting in the winter. The patient also has the flexibility of scheduling their follow-up eVisit at a time and location convenient for them and their family. Family members also experience a benefit by being able to join the eVisit remotely, offering increased support to patients, which is particularly crucial for seniors. Another value to the eVisit context was the ability to complete an accurate medication reconciliation as patients always had access to their actual medications at home, and they could show the real prescription with labels, and these were tallied with the medication list obtained from the patient's pharmacy before the eVisit.

Population Health

Perhaps one of the most significant outcomes we report is the reduced patient wait time-to follow-up. eVisits allowed the physician to see patients sooner than would be possible for an in-person encounter, thus increasing the availability of health care. The technology supporting eVisits also provided the ability to share imaging or echocardiographic findings with the patient in addition to sharing medical illustrations, enabling and facilitating patient education, understanding, and empowerment. eVisits also offer the ability to identify risks and patient vulnerabilities sooner, improve treatment adherence, and support behavioral and care interventions to improve speech, mobility, and enhance access to home care or community-based care or allied health services. Flexible scheduling allows physicians to be more productive with their time, enabling them to distribute their clinical activity to accommodate other commitments including teaching, research, and administration. In addition to increased productivity, using eVisit has the potential to address some of the significant contributors to physician burnout (including work and organizational factors), which can in turn have consequences on patient care and health care costs [39].

Reducing Per Capita Cost

Reducing Out-of-Pocket Patient Cost

The eVisit offers the potential for a significant reduction in per capita costs for outpatient care. The estimated direct OOP cost savings for a single in-person visit is considerable. This could be much higher if accounting for multiple health care encounters. Our estimated OOP cost saving per visit is probably conservative; real savings would vary significantly based on the individual's hourly wage, employment status, other personal factors, and visit characteristics.

Opportunity Cost to Health Care

Opportunity cost refers to the cost or money that the health care system could have allocated or used for similar or different interventions [19]. The estimates of the opportunity cost of using in-person care for our study cohort are Can $23,832 to $28,584. The opportunity cost could represent an opportunity to provide outpatient in-person health care to a different segment of the population. Thus, with the same health care budget, health care services could be offered to a higher number of patients, driving down per capita health care expenditure. Alternatively, this opportunity cost could be redirected to other high priority areas to increase health care outcomes and efficiency. Scaling of the project within health care organizations could have a significant impact.

Limitations of the Electronic Visit

Some of the disadvantages of the eVisits relate to the technology itself. The service cannot be offered to patients who do not have an internet enabled device and/or sufficient internet connection speed, thus potentially limiting the access to home-based eVisits to patients with lower socioeconomic status. In addition, internet access and speed are limited in some geographic areas, especially in rural and remote communities. Moreover, patients with physical, cognitive, and language disabilities may find it hard to use the technology or navigate the appropriate software on their own. Another disadvantage is the inability to do an in-person clinical examination, limiting the utility of the eVisits in some clinical scenarios. However, the video-based examination has been shown to be reliable and valid [40,41].

Strengths of This Project

The strengths of the pilot study include implementation of a successful eVisit program for outpatient follow-up in a specialty stroke clinic catering largely to a senior population. The combination of high degree of patient satisfaction with potential savings of both time and money holds promise for economically improving access to care. The mean time spent per eVisit, including the documentation of the clinical encounter, was 10 min, which is comparable with the time allocated for the physician-patient portion of a typical in-person clinical contact. We believe that our choice to perform a test eVisit before the physician-patient visit reduced the chance of communication technical difficulty and resulted in very successful physician eVisits with 5% (4/81) failure rate and 3% (2/81) no-show rate. This pilot study will inform the expansion of the eVisit project to other specialty clinics within the organization as the next phase.

Limitations of This Study

Some limitations of the study include the limited sample size and pilot duration. The scope of the project was narrow, involving 1 specialty clinic. These limitations prevent broader generalizability. There is a potential for bias in assessing patient satisfaction as a result of using a brief telephone survey with nonprobability sampling. The economic analysis of outpatient costs, as well as opportunity cost for in-person care, needs to be corroborated in future studies across multiple organizations. The outpatient health care costing data may vary amongst different health care organizations.

Future Directions

The limited uptake of virtual care services such as eVisits by physicians was recently reported in a 2018 survey, highlighting the need for appropriate reimbursement or alternative payment models as well as improved technology, privacy and security guidelines, and support from clinician associations and governance bodies [32]. Physician eVisits have recently been approved for reimbursement by the MOHLTC, Ontario, as part of a pilot project through OTN, which may increase utilization. However, technology integration and improvements, workflow, and processes or quality improvements, and privacy and security guidelines are required to integrate eVisits into the health care system seamlessly. In our pilot study, the majority of patients used a tablet or mobile phone, likely because of the widespread ownership of such devices and ease of use with internet-enabled smart devices even among seniors in Canada. This raises the potential to develop and integrate more mobile-based telehealth solutions. On the basis of the acceptance and efficacy of the pilot study, there has been a 40.2% (76/189) conversion of follow-ups from in-person to eVisit in the stroke prevention clinic pilot catering to seniors.

To address some of the limitations of eVisit, *eVisit stations* could be established at community health care centers, where local staff could be easily trained to help set up an eVisit. This could allow patients living in rural or remote areas, without access to the internet or who are unable to independently use the eVisit system, to attend an appointment locally and avoid lengthy travel to an urban center. A more comprehensive health economic analysis of eVisits for outpatient care is warranted in any future studies given the potential for reduction in per-capita health care costs. In addition, the impact of the eVisits on patient outcomes, wait times, readmission rates, and impact on the use of urgent care or emergency departments should be evaluated.

Conclusions

eVisits were implemented successfully for an outpatient follow-up clinic for adult stroke patients in our pilot study. eVisits were well received by patients and consistent with a patient-centered care philosophy. eVisits have the potential to significantly transform the ambulatory clinic practice by addressing some of the barriers to care and improving patient experience, reducing per capita health care costs, and improving population health. Such a transformative change needs the involvement of health care professionals, health services researchers or economists, hospital leadership, clinician associations, and health system governance bodies at the regional and provincial levels to inform evidence-based practice guidelines and sustainable models of care. eVisits are scalable and could be expanded to additional specialty programs, a move that is underway at KHSC.

Acknowledgments

The authors would like to thank Ms Maria Foss for her Secretarial assistance in coordinating the eVisits in the stroke prevention clinic.

This eVisit pilot project was supported by funding from the Department of Medicine, Queen’s University. The funding was required primarily for part-time secretarial salary support.

Conflicts of Interest

RA received internal grant funding for the eVisit pilot project from the Department of Medicine, Queen’s University.

Multimedia Appendix 1
Definitions used for Time, Distance and Cost savings.
[PDF File (Adobe PDF File), 17 KB - jmir_v21i10e13734_app1.pdf]

Multimedia Appendix 2
Volume of patients seen in the stroke prevention clinic during the pilot phase.
[PDF File (Adobe PDF File), 24 KB - jmir_v21i10e13734_app2.pdf]

Multimedia Appendix 3
Distance and time savings for patients.
[PDF File (Adobe PDF File), 15 KB - jmir_v21i10e13734_app3.pdf]

Multimedia Appendix 4
Ambulatory care costs for CACS codes (E751 and E752) in Ontario and Alberta for 2016/2017.

[PDF File (Adobe PDF File), 27 KB - jmir_v21i10e13734_app4.pdf]

References

1. Sibley LM, Glazier RH. Reasons for self-reported unmet healthcare needs in Canada: a population-based provincial comparison. Healthc Policy 2009 Aug;5(1):87-101 [FREE Full text] [Medline: 20676253]
2. Ronksley PE, Sanmartin C, Campbell DJ, Weaver RG, Allan GM, McBrien KA, et al. Perceived barriers to primary care among western Canadians with chronic conditions. Health Rep 2014 Apr;25(4):3-10 [FREE Full text] [Medline: 24744042]
3. Kasman NM, Badley EM. Beyond access: WHO reports that health care is not being received when needed in a publicly-funded health care system? Can J Public Health 2004;95(4):304-308. [Medline: 15362477]
4. Gibbard R. Canadian Medical Association. 2018. Meeting the Care Needs of Canada's Aging Population URL: http://live-cma-ca.pantheonsite.io/sites/default/files/pdf/Activities/Conference%20Board%20of%20Canada%20-%20Meeting%20the%20Care%20Needs%20of%20Canada's%20Aging%20Population.PDF; (Archived by WebCite® at http://www.webcitation.org/75cwQ2psA) [accessed 2019-01-24]
5. Canadian Institute for Health Information. 2011. Health Care in Canada, 2011 - A Focus on Seniors and Aging URL: https://secure.cihi.ca/free_products/HCIC_2011_seniors_report_en.pdf [accessed 2019-01-23] [WebCite Cache ID 75owqBcHw]
6. Turcotte M, Schellenberg G. Statistics Canada. 2006. A Portrait of Seniors in Canada: Introduction URL: https://www150.statcan.gc.ca/n1/pub/89-519-x/89-519-x2006001-eng.htm [accessed 2019-02-14] [WebCite Cache ID 76B4InSQj]
7. Picker: Patient Centred Care Improvement. 2018. Principles of Person Centred Care URL: https://www.picker.org/about-us/picker-principles-of-person-centred-care/ [accessed 2019-01-22] [WebCite Cache ID 75cwYdfIA]
8. Stiefel M, Nolan K. Institute for Healthcare Improvement: IHI. 2012. A Guide to Measuring the Triple Aim: Population Health, Experience of Care, and Per Capita Cost URL: http://www.ihi.org/resources/Pages/IHIWhitePapers/AGuidetoMeasuringTripleAim.aspx [accessed 2019-01-24] [WebCite Cache ID 75frGAdU9]
9. Canadian Institute for Health Information. 2018. How Canada Compares: Results From The Commonwealth Fund's 2017 International Health Policy Survey of Seniors URL: https://secure.cihi.ca/estore/productFamily.htm?pf=PFC3650&lang=en&media=0 [accessed 2019-01-24] [WebCite Cache ID 75dun4JBr]
10. Uppal S, Barayandema A. Statistics Canada. 2016. Life Satisfaction Among Canadian Seniors URL: https://www150.statcan.gc.ca/n1/pub/75-006-x/2018001/article/54977-eng.htm [accessed 2019-01-22] [WebCite Cache ID 75cvf9gO4]
11. Canada Health Infoway: Digital Health in Canada. 2019. Showing Results for Resources and Web Pages Tagged: Virtual Visits URL: https://www.infoway-inforoute.ca/en/solutions/access-health/access-to-services/virtual-visits [accessed 2019-05-28] [WebCite Cache ID 78i7zKnog]
12. McGrail KM, Ahuja MA, Leaver CA. Virtual visits and patient-centered care: results of a patient survey and observational study. J Med Internet Res 2017 May 26;19(5):e177 [FREE Full text] [doi: 10.2196/jmir.7374] [Medline: 28550006]
13. Canada Health Infoway: Digital Health in Canada. 2018. Connecting Patients for Better Health: 2018 URL: https://www.infoway-inforoute.ca/en/component/edocman/3564-connecting-patients-for-better-health-2018/view-document?Itemid=0; [accessed 2019-01-24] [WebCite Cache ID 75ctcI4dZ]
14. Accenture. 2017. Virtual Care: Just What the Consumer Ordered URL: https://www.accenture.com/t20181011T170958Z__w__/us-en/_acnmedia/PDF-88/Accenture-Strategy-DD-Virtual-Care-POV.pdf [accessed 2019-01-23] [WebCite Cache ID 75duyw71q]
15. Bhatia RS, Falk W. CD Howe Institute. 2018. Modernizing Canada's Healthcare System through the Virtualization of Services URL: http://www.ams-inc.on.ca/wp-content/uploads/2015/12/Modernizing-Canadas-Healthcare-System-through-the-Virtualization-of-Services-by-Will-Falk.pdf [accessed 2019-01-24] [WebCite Cache ID 75duV5bdZ]
16. Duffy S, Lee TH. In-person health care as option B. N Engl J Med 2018 Jan 11;378(2):104-106. [doi: 10.1056/NEJMp1710735] [Medline: 29320653]
17. Coles N, Patel BP, Li P, Cordeiro K, Steinberg A, Zdravkovic A, et al. Breaking barriers: adjunctive use of the Ontario telemedicine network (OTN) to reach adolescents with obesity living in remote locations. J Telemed Telecare 2018 Dec 10:- (epub ahead of print). [doi: 10.1177/1357633X18816254] [Medline: 30526258]
18. du Plessis V, Roland B, Bollman RD. Statistics Canada. 2002. Definitions of 'Rural' URL: https://www150.statcan.gc.ca/n1/pub/21-601-m/21-601-m2002061-eng.htm [accessed 2019-02-14] [WebCite Cache ID 76B3VBxlp]
19. Palmer S, Raftery J. Economic notes: opportunity cost. Br Med J 1999 Jun 5;318(7197):1551-1552 [FREE Full text] [doi: 10.1136/bmj.318.7197.1551] [Medline: 10356019]
20. Statistics Canada. 2018. Average Usual Hours and Wages by Selected Characteristics, Monthly, Unadjusted for Seasonality, Last 5 Months (X 1,000) URL: https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=1410032002&pickMembers%5B0%5D=3.4; [accessed 2019-01-23] [WebCite Cache ID 75eChheNc]
21. Canadian Institute for Health Information. 2018. Comprehensive Ambulatory Classification System (CACS) URL: https://www.cihi.ca/en/comprehensive-ambulatory-classification-system-cacs [accessed 2019-02-06]
22. Alberta Wait Times Reporting: Wait Time Trends. 2019. Health Costing, Interactive Health Data Application URL: http://www.ahw.gov.ab.ca/IHDA_Retrieval/selectCategory.do [accessed 2019-02-06]

23. Government of Ontario. 2019. Ontario Case Costing Initiative (OCCI) URL: https://hsim.health.gov.on.ca/hdbportal/ [accessed 2019-02-06]
24. Sirintrapun SJ, Lopez AM. Telemedicine in cancer care. Am Soc Clin Oncol Educ Book 2018 May 23;38:540-545. [doi: 10.1200/EDBK_200141] [Medline: 30231354]
25. Patel UK, Malik P, DeMasi M, Lunagariya A, Jani VB. Multidisciplinary approach and outcomes of tele-neurology: a review. Cureus 2019 Apr 8;11(4):e4410 [FREE Full text] [doi: 10.7759/cureus.4410] [Medline: 31205830]
26. Lin LA, Casteel D, Shigekawa E, Weyrich MS, Roby DH, McMenamin SB. Telemedicine-delivered treatment interventions for substance use disorders: a systematic review. J Subst Abuse Treat 2019 Jun;101:38-49. [doi: 10.1016/j.jsat.2019.03.007] [Medline: 31006553]
27. Kane-Gill SL, Rincon F. Expansion of telemedicine services: telepharmacy, telestroke, teledialysis, tele-emergency medicine. Crit Care Clin 2019 Jul;35(3):519-533. [doi: 10.1016/j.ccc.2019.02.007] [Medline: 31076051]
28. Galea MD. Telemedicine in rehabilitation. Phys Med Rehabil Clin N Am 2019 May;30(2):473-483. [doi: 10.1016/j.pmr.2018.12.002] [Medline: 30954160]
29. Mehrotra A, Uscher-Pines L, Lee MS. The dawn of direct-to-consumer telehealth. In: Rheuban KS, Krupinski EA, editors. Understanding Telehealth. New York: McGraw-Hill Education; 2018:217-223.
30. Dare F. Accenture. 2017. Voting for Virtual Healthcare: Virtual Healthcare for the 21st Century URL: https://www.accenture.com/us-en/insight-voting-virtual-health-survey [accessed 2018-12-14] [WebCite Cache ID 74fBs03Gu]
31. Paré G, Leaver C, Bourget C. Diffusion of the digital health self-tracking movement in Canada: results of a national survey. J Med Internet Res 2018 May 2;20(5):e177 [FREE Full text] [doi: 10.2196/jmir.9388] [Medline: 29720359]
32. Canada Health Infoway: Digital Health in Canada. 2018. 2018 Canadian Physician Survey: Physicians' Use of Digital Health and Information Technologies in Practice URL: https://www.infoway-inforoute.ca/en/component/edocman/3643-2018-canadian-physician-survey/view-document?Itemid=101; [accessed 2019-01-22] [WebCite Cache ID 75ctqRhz8]
33. Elliott T, Shih J. Direct to consumer telemedicine. Curr Allergy Asthma Rep 2019 Jan 19;19(1):1. [doi: 10.1007/s11882-019-0837-7] [Medline: 30661123]
34. Seritan AL, Heiry M, Iosif AM, Dodge M, Ostrem JL. Telepsychiatry for patients with movement disorders: a feasibility and patient satisfaction study. J Clin Mov Disord 2019;6:1 [FREE Full text] [doi: 10.1186/s40734-019-0077-y] [Medline: 31183157]
35. Korn RE, Shukla AW, Katz M, Keenan HT, Goldenthal S, Auinger P, et al. Virtual visits for Parkinson disease: a multicenter noncontrolled cohort. Neurol Clin Pract 2017 Aug;7(4):283-295 [FREE Full text] [doi: 10.1212/CPJ.0000000000000371] [Medline: 28840919]
36. Hanson RE, Truesdell M, Stebbins GT, Weathers AL, Goetz CG. Telemedicine vs office visits in a movement disorders clinic: comparative satisfaction of physicians and patients. Mov Disord Clin Pract 2019 Jan;6(1):65-69. [doi: 10.1002/mdc3.12703] [Medline: 30746418]
37. Finkelstein JB, Cahill D, Kurtz MP, Campbell J, Schumann C, Varda BK, et al. The use of telemedicine for the postoperative urological care of children: results of a pilot program. J Urol 2019 Jul;202(1):159-163. [doi: 10.1097/JU.0000000000000109] [Medline: 30707132]
38. Carotenuto A, Rea R, Traini E, Ricci G, Fasanaro AM, Amenta F. Cognitive assessment of patients with Alzheimer's disease by telemedicine: pilot study. JMIR Ment Health 2018 May 11;5(2):e31 [FREE Full text] [doi: 10.2196/mental.8097] [Medline: 29752254]
39. Patel RS, Bachu R, Adikey A, Malik M, Shah M. Factors related to physician burnout and its consequences: a review. Behav Sci (Basel) 2018 Oct 25;8(11):E98 [FREE Full text] [doi: 10.3390/bs8110098] [Medline: 30366419]
40. Craig JJ, McConville JP, Patterson VH, Wootton R. Neurological examination is possible using telemedicine. J Telemed Telecare 1999;5(3):177-181. [doi: 10.1258/1357633991933594] [Medline: 10628033]
41. Bove R, Bevan C, Crabtree E, Zhao C, Gomez R, Garcha P, et al. Toward a low-cost, in-home, telemedicine-enabled assessment of disability in multiple sclerosis. Mult Scler 2018 Aug 24:- (epub ahead of print). [doi: 10.1177/1352458518793527] [Medline: 30141729]

Abbreviations

CACS: Comprehensive Ambulatory Classification System
CIHI: Canadian Institute for Health Information
eVisit: electronic visit
IQR: interquartile range
KHSC: Kingston Health Sciences Centre
MOHLTC: Ministry of Health and Long-Term Care
OOP: out-of-pocket
OTN: Ontario Telemedicine Network
SDM: substitute decision maker

Edited by G Eysenbach, N Cahill, S Hagens; submitted 26.02.19; peer-reviewed by S Bhatia, I Mircheva; comments to author 18.05.19; revised version received 13.07.19; accepted 21.08.19; published 07.10.19.

Please cite as:
Appireddy R, Khan S, Leaver C, Martin C, Jin A, Durafourt BA, Archer SL
Home Virtual Visits for Outpatient Follow-Up Stroke Care: Cross-Sectional Study
J Med Internet Res 2019;21(10):e13734
URL: https://www.jmir.org/2019/10/e13734
doi:10.2196/13734
PMID:31593536

Original Paper

Hospital-Based Back Surgery: Geospatial-Temporal, Explanatory, and Predictive Models

Lawrence Fulton[1], PhD; Clemens Scott Kruse[1], PhD

Department of Health Administration, Texas State University, San Marcos, United States

Corresponding Author:
Lawrence Fulton, PhD
Department of Health Administration
Texas State University
Encino Hall
712 N Comanche St
San Marcos, 78666
United States
Phone: 1 (512) 245 3492
Email: lf25@txstate.edu

Abstract

Background: Hospital-based back surgery in the United States increased by 60% from January 2012 to December 2017, yet the supply of neurosurgeons remained relatively constant. During this time, adult obesity grew by 5%.

Objective: This study aimed to evaluate the demand and associated costs for hospital-based back surgery by geolocation over time to evaluate provider practice variation. The study then leveraged hierarchical time series to generate tight demand forecasts on an unobserved test set. Finally, explanatory financial, technical, workload, geographical, and temporal factors as well as state-level obesity rates were investigated as predictors for the demand for hospital-based back surgery.

Methods: Hospital data from January 2012 to December 2017 were used to generate geospatial-temporal maps and a video of the Current Procedural Terminology codes beginning with the digit 63 claims. Hierarchical time series modeling provided forecasts for each state, the census regions, and the nation for an unobserved test set and then again for the out-years of 2018 and 2019. Stepwise regression, lasso regression, ridge regression, elastic net, and gradient-boosted random forests were built on a training set and evaluated on a test set to evaluate variables important to explaining the demand for hospital-based back surgery.

Results: Widespread, unexplained practice variation over time was seen using geographical information systems (GIS) multimedia mapping. Hierarchical time series provided accurate forecasts on a blind dataset and suggested a 6.52% (from 497,325 procedures in 2017 to 529,777 in 2018) growth of hospital-based back surgery in 2018 (529,777 and up to 13.00% by 2019 [from 497,325 procedures in 2017 to 563,023 procedures in 2019]). The increase in payments by 2019 are estimated to be US $323.9 million. Extreme gradient-boosted random forests beat constrained and unconstrained regression models on a 20% unobserved test set and suggested that obesity is one of the most important factors in explaining the increase in demand for hospital-based back surgery.

Conclusions: Practice variation and obesity are factors to consider when estimating demand for hospital-based back surgery. Federal, state, and local planners should evaluate demand-side and supply-side interventions for this emerging problem.

(J Med Internet Res 2019;21(10):e14609) doi:10.2196/14609

KEYWORDS

back surgery; neurosurgeon; elastic net; lasso; ridge; random forest; geospatial mapping; health economics; obesity; practice variation

Introduction

Background

In 2012, there were 3689 practicing board-certified neurosurgeons in the United States [1]. That number was largely unchanged in 2016 [2]. During these years, demand for back surgery (Current Procedural Terminology [CPT] codes beginning with the digit 63) increased by 49% from 311,028 to 464,391, and by the end of 2017, that increase was 60% [3]. CPT *63* medical codes are a series of spinal procedures including laminectomies, laminotomies, decompressions, and

corpectomies. These procedures do not include needle decompression, catheter implantation, and, as of 2019, endoscopic decompression [4]. Given the stable supply and increasing demand, it is not surprising that the average payment procedure increased from US $4166 to US $4859 from 2012 to 2016 and to US $5452 by the middle of 2018, an effective 4.5% inflation rate [3]. Forecasting models that address increasing demand are necessary to evaluate potential supply and demand-side interventions.

Unsurprisingly, there is a marked variation in the treatment of back disorders such as spondylolisthesis [5]. This variation affects costs [6] as well as outcomes [5] associated with back surgery. The implication of this variation is increased demand. By evaluating the current geographic demand, policy makers can prioritize efforts for cost and variation reduction by evaluating those states and counties that exhibit high practice area variation, implementing evidence-based best practice policies and guidelines, educating populations about obesity risks, and implementing interventions for those at risk of obesity (eg, those living in food deserts).

During the same time that back surgeries have increased, adult obesity rates in the United States have also increased. The rate of this increase was 5% from 2012 to 2016 (34.9%-39.6%) [7]. Obesity has been linked to increased costs of medical care [8]. Although obese patients benefit from at least some back surgeries, they do not fare as well as nonobese patients [9]. Although obesity has been linked to back pain [10], no studies were found that directly link obesity to back surgery requirements. This study evaluated that relationship as well.

Objectives

This study addressed 3 specific aspects of hospital-based CPT 63 surgery. First, a geospatial-temporal analysis by zip code is conducted to describe the previous and current demand for CPT 63 surgery. The significance of this geospatial-temporal analysis is that practice variation is highlighted for evaluation by federal, state, and local policy makers. Second, forecasting models estimate the demand and payments overall, by census region and by state. These models are also designed for state policy makers to assess potential supply- and demand-side intervention requirements. Third, explanatory models are developed to correlate obesity rates and financial, technical, workload, temporal, and geospatial variables with demand for CPT 63 procedures. This analysis does not appear to be previously investigated and is an important but overlooked correlational analysis. The study focused specifically on hospital-based knee surgery with CPT 63 codes (some of which reflect inpatient procedures) and was delimited to knee surgery only.

Methods

Data

Definitive Healthcare provided the hospital, zip code, and state-level procedure and cost data from January 2012 to June 2018 through the hospital *revenue center analytics* query, which includes queries by CPT code. Data in Definitive Healthcare are derived from the Standard Analytical Files by the Centers for Medicare and Medicaid Services (CMS)]. From these data, the organization uses undisclosed algorithms to estimate all-payor claims. Columns with fewer than 11 claims or procedures are not shown because of privacy requirements [3]. For this analysis, only complete annual data from 2012 to 2017 were used, as the CMS datafile and thus the associated estimates from January 2018 through June 2018 were approximately only 93% complete [3].

The Centers for Disease Control and Prevention's Behavioral Risk Factor Surveillance System (BRFSS) prevalence data provided the information for state-level adult obesity rates by year, from 2012 to 2017 [11]. Guam, Puerto Rico, and the US Virgin Islands were excluded from the analysis because of small sample sizes in both datasets.

Geospatial Analysis

Heat maps are used to plot the zip code unit of analysis procedure data by year. Heat maps provide the intensity of the number of claims by time and geographic region. These types of maps have been used for improving minority health surveillance [12], examining birth outcomes [13], and evaluating a variety of other applications in health care. The value in geospatial-temporal analysis is the graphical depiction of change in demand over time. A video display from 2012 to 2017 with standardized heat intensities provides an animated view of the change in demand by location. An analysis of cost and demand centers is then provided.

Forecasting Analysis

The data in the Definitive Healthcare dataset are nonseasonal as they provide annual-level observations by the hospital unit of analysis. Even so, generating nonseasonal forecasting models that have predictive capability on a blind withhold set at the proper level of aggregation can provide decision support for supply- and demand-side interventions. These types of models have found support in many areas of health care such as radiology [14] and Alzheimer disease [15].

To this end, hierarchical time series (HTS) [16] using R statistical software [17] evaluated the number of claims as a function of time series components. An HTS recognizes that data are aggregated at various levels. In this case, the hierarchy evaluated include the states, the census regions, and the nation. The models are built on a training set of data for the years 2012 to 2015 and forecast on a blind test set, years 2016 and 2017. Although Bayesian hierarchical models have been used for spatially correlated health outcomes and utilization rates [18], there is no readily found use of HTS for prediction in health care.

To understand HTS, one needs to only consider a single medical system that operates in 2 separate states with 3 hospitals per state. There are then 4 basic ways using which one might forecast annual visits as an example:

1. A forecast might be generated for each hospital, aggregated at the state level and then further aggregated at the system level. A variety of different forecast methods might be used to generate the forecasts. The term for this method is bottoms up.

2. Forecasts might be generated at the state level and then disaggregated (eg, via historical proportions) to the hospitals and aggregated to the system level. Again, the forecasts might be generated in multiple ways. The term for this method is middle out.
3. Forecasts might be generated at the system level (via multiple methods) and then disaggregated to all levels below (eg, proportions). The term for this is top down.
4. One might take some combination of the previous methods to minimize forecast error. This is an ensemble method that might be termed the optimal reconciliation approach, which is optimal if the forecasts are unbiased [19].

To avoid selection bias, all methods were evaluated for performance on the test set. Furthermore, the method for forecasting at these levels of hierarchy was using autoregressive integrated moving average (ARIMA) components and as well as smoothed error and trend components (exponential, trend, seasonality [ETS] without seasonality).

ARIMA models focus on autocorrelation of components for stationary time series data. The *AR* components are autoregressive terms, offset by time. For example, the number of claims at time *t* might be forecast by using the number of claims at time *t-1*. This would be an AR1 model, as there is 1 offset. As ARIMA models assume stationarity of the time series for forecasting, 2 other components are necessary. The first is differencing or *integration*, the *I* in ARIMA, which helps stabilize the mean (whereas transformations help stabilize the variance). Although seasonality and trend might make an ARIMA nonstationary, differencing often corrects this. Sometimes, more than 1 difference is required to make the time series stationary, for example, y_t-y_{t-1}-(y_{t-1}-y_{t-2}) is a 2d order differencing. The last component, the *MA* or moving average, corrects for autocorrelated errors as well. This component averages previous observation(s) with the previous forecast(s) [19].

ETS models have 3 components: error, trend, and seasonality. As the data in this study are not seasonal, only the error smoothing (identical to a moving average) and the trend component (a Holt model [20]) are evaluated.

With HTS bottom-up models, a separate ARIMA/ETS is built for each bottom-level component. For middle-out models, all middle-level components have separate forecasts. For top-down models, a single forecast is generated and proportioned down to the lower levels.

Explanatory Analysis

Stepwise linear regression (both forward and backward), lasso regression, robust regression, elastic net regression, and extreme gradient-boosted random forests are built on unaggregated data as well as state-level aggregated data to estimate the number of claims. These models are built on a random 80% training set (10,771 unaggregated, 245 aggregated observations) and evaluated on a 20% withhold set (2693 unaggregated, 61 aggregated) as well. The total number of valid observations were 13,464 unaggregated and 306 aggregated. The primary hypothesis is that the inclusion of obesity rates as an independent variable will yield better explanatory models for both the number of claims and the payment per claim.

Stepwise linear regression based on minimum Akaike Information Criterion was selected over best subset because of the computational complexity. By using forward and backward simultaneously, variables are added in sequence but might be removed if they no longer contribute to the model [21].

Lasso regression is a form of constrained regression that penalizes a model that selects too many variables by using an L1-norm formulation (absolute value), whereas ridge regression is similar but penalizes using an L2-norm formulation (squared coefficient estimates). Elastic net uses a weighted L1 and L2 norm penalty function to reduce the number of coefficients in the model. Formulae for estimating the parameters of the linear model, the lasso regression, the ridge regression, and the elastic net are shown in Figure 1.

Figure 1. Argmin equations for the regression models.

Linear regression (OLS):
$$\hat{\beta} = \sum_{i=1}^{N}(y_i - \beta_0 - \sum_{j=1}^{p} x_{ij}\beta_j)^2$$

Lasso regression (L1-norm):
$$\hat{\beta} = \sum_{i=1}^{N}(y_i - \beta_0 - \sum_{j=1}^{p} x_{ij}\beta_j)^2 + \lambda \sum_{j=1}^{p} |\beta_j|$$

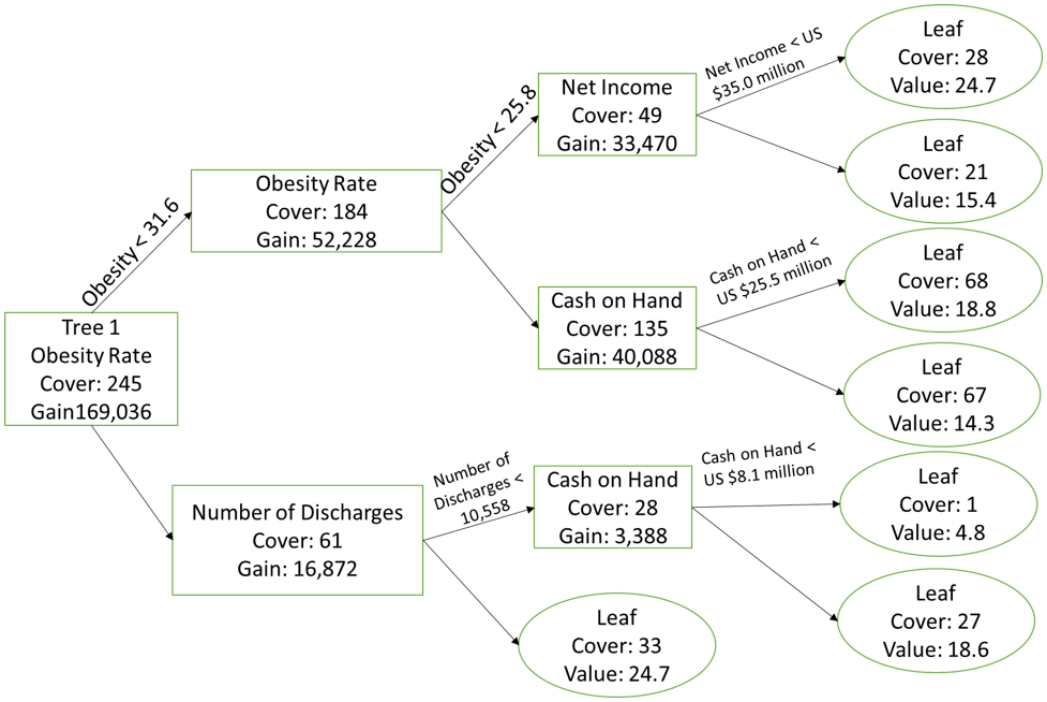
Ridge regression (L2-norm):
$$\hat{\beta} = \sum_{i=1}^{N}(y_i - \beta_0 - \sum_{j=1}^{p} x_{ij}\beta_j)^2 + \lambda \sum_{j=1}^{p} \beta_j^2$$

Elastic net (L1 and L2 norm):
$$\hat{\beta} = \sum_{i=1}^{N}(y_i - \beta_0 - \sum_{j=1}^{p} x_{ij}\beta_j)^2 + \lambda \sum_{j=1}^{p} (\alpha\beta_j^2 + (1-\alpha)|\beta_j|)$$

Random forests, a machine learning technique, use an ensemble of decorrelated tree models and average the estimates of those trees to build forecasts. A tree model itself classifies counts of observations by splitting variables at points based on some

decision criteria. An example of a tree with a depth of 3 (3 branches) is given in Figure 2, which splits observations by obesity rate less than 31.63 and then again by obesity rate less than 25.75 and number of discharges less than 10,558.78 and then finally by net income less than US $35,018,392, cash less than US $25,522,424, and cash less than US $8,122,498 [21]. The graph was produced by the xgboost package of R [22]. Gradient-boosted random forests optimize a cost function based on the (pseudo-)residuals of a function using nonlinear optimization techniques. Essentially, the residuals of each tree in the forest are refitted with the possible independent variables in another tree model to estimate a better fit of the original function. Often, a learning rate (shrinkage) is applied to the residuals to allow for better generalization. A discussion of gradient boosting is available in Chapter 10 of *Elements of Statistical Learning* [21].

Figure 2. An example of a single tree model with 3 branches. The graph was produced by the xgboost package of R. (NumDischarges indicates the number of discharges).



Variables

All the considered variables from the Definitive Healthcare dataset are shown in Table 1 with reasons for exclusion/inclusion. Most variables were, by default, included during analysis; however, those variables that were linear combinations of each other or were necessarily unknowable when forecasting CPT 63 codes were omitted.

There is 1 primary dependent variable of interest: number of claims for CPT 63 codes. This variable is measured at the hospital level over time and is also aggregated by zip code/year for geospatial mapping and by state/year for forecasting and additional modeling analysis. The number of claims include third-party invoices provided by the hospital, regardless of the payer. The number of claims provides a measure of the met demand for services.

For the geospatial and temporal analyses, the variables year and zip code (aggregated hospital-level data) are used to describe the intensity of both the number of claims and the payment per claim. Zip code provides a high resolution for geographic claims data. For the HTS forecasting analysis, time components are used without external regressors.

Explanatory stepwise regression, lasso regression, ridge regression, elastic net, and gradient-boosted random forest models investigate financial variables, technical variables, workload variables, geospatial variables, a temporal variable (year), and obesity rates (defined as the proportion of individuals with a body mass index greater than or equal to 30%). A discussion of each of the variable groups and variables follows.

The financial variables investigated include net patient revenue, net income, cash on hand, total assets, total liabilities, and proportion of Medicare/Medicaid reimbursement. The financial variables were carefully selected from the set of available financials such that they are not a linear combination of other variables or nearly a linear combination (see Table 1). Although available, total payments and charges for CPT 63 were not used in the models, as they (1) would not be known in advance and (2) would necessarily be direct functions of the number of claims.

Quantitative workload variables include the number of staffed beds, discharges, emergency room visits, surgeries, affiliated physicians, and employees. Categorical technical variables include ownership type, medical school affiliation, and hospital type. These variables are investigated because of their availability and possible confounding effects. Geographic variables include urban/rural location, state, and zip code. These variables are important in evaluating practice area variation and associated effects.

Obesity rates are of interest to the study. These rates are assigned based on the state, as county and zip code level data are not available. This independent variable is of importance to the study.

Table 1. Variables in the study.

Variable	Type	Definition	Scale of measurement
Number of claims	Dependent	Filed third-party claims	{0, 1, 2, ...k}
Obesity rate	Obesity	Percentage obese by state	[0%, 40%]
CPT[a] 63 payments	Financial	Total CPT 63 payments	US $
CPT 63 charges	Financial	Total CPT 63 charges	US $
Net patient revenue	Financial	Total revenue from patients	US $
Total revenues	Financial	All revenue, patient-related or otherwise	US $
Net income	Financial	Revenues less expenses	US $
Total expenses	Financial	Total dollars attributed to expenses	US $
Cash on hand	Financial	Funds immediately available	US $
Total assets	Financial	Current and noncurrent assets	US $
Total liabilities	Financial	Current and long-term debt	US $
Percentage Medicare/Medicaid	Financial	Percentage claims from either source	%
State	Geospatial	Hospital's state (address)	AK, AL, ...
Zip code	Geospatial	5-digit hospital zip code	78666, ...
Geographic classification	Technical	Rural or urban location	Rural, urban
Ownership	Technical	Hospital ownership	Nonprofit, profit, government
Medical school affiliation	Technical	Level of affiliation if any	Graduate, major, limited, none
Hospital type	Technical	Type of hospital	Short-term acute, children's, etc
Year	Temporal	Year of report	2012, 2018
Number of staffed beds	Workload	Per Medicare report	{0, 1, ...n}
Number of discharges	Workload	Total number of inpatient discharges	{0, 1, ...n}
Number of Medicare discharges	Workload	Number of Medicare discharges	{0, 1, ...n}
Estimated number of emergency room visits	Workload	Number of emergency room visits	{0, 1, ...n}
Total surgeries	Workload	Number of surgeries	{0, 1, ...n}
Total acute days	Workload	Number of acute bed days	{0, 1, ...n}
Number of affiliated physicians	Workload	Number of affiliated physicians	{0, 1, ...n}
Number of employees	Workload	Number of employees	{0, 1, ...n}

[a]CPT: Current Procedural Terminology.

Results

Descriptive Statistics: Missing Data

Missing data were present in the Definitive Healthcare dataset. As the percentage of missing data was small, the data were imputed via regression trees (simple imputation). The total number of valid observations at the hospital unit of analysis from January 2012 to December 2017 was 13,769. There were 2244 unique zip codes with data resulting in 13,464 observations from 2012 to 2017, although many of these were true zeros. Aggregated at the state level, there were 306 observations of the 50 states plus the District of Columbia over the 6-year span.

Descriptive Statistics: Quantitative Data

Important descriptive statistics for the Definitive Healthcare data are shown (Table 2). The average number of CPT 63 claims by hospital by year was 182, and the average payment was over US $4045.99, about 50.37% of mean charges (US $8,032.13). On average, hospitals performing these claims were large (227 beds with 1629 employees and 299 affiliated physicians). These hospitals had on average positive net income (US $22 million) and assets exceeding liabilities. Overall, 45% of their patients used Medicare or Medicaid reimbursement.

Table 2. Descriptive statistics for all years and all hospitals (N=13,769). The "K" suffix indicates dollars in thousands, while the "M" suffix indicates dollars in millions.

Variable	Mean (SD)	Median	Minimum	Maximum
Claims, n	182 (244)	89	11	3592
Payments/claim, US $	4045.99 (2448.96)	3659.12	0	34,975.76
Payments, US $	767.9K (1097.8K)	366.2K	0	18,966.3K
Charges, US $	1517.6K (2467.4K)	616.4K	0	35,317.2K
Charges/claim, US $	8032.13 (6993.90)	6370.08	0	137,058.80
Net patient revenue, US $	343.8M (430.4M)	217.5M	−98.6M	5340.9M
Net income, US $	22.8M (102.8M)	11.3M	−1648M	1316.0M
Cash, US $	30.3M (145.3M)	2.8M	−1992.7M	3597.8M
Total assets, US $	443.0M (820.1M)	203.5M	−231.7M	9969.4M
Total liabilities, US $	178.2M (465.7M)	69.2M	−2583.8M	6372.4M
Staffed beds, n	227.45 (201.36)	177.00	1.00	2626
Discharges, n	11,822.61 (11,395.52)	8899.00	1.00	127,600
Emergency room visits, n	47,439.59 (39,580.09)	39,209.00	0	543,457
Surgeries, n	9643.21 (9666.56)	7019.00	0	134,638
Affiliated physicians, n	298.57 (333.43)	198.00	1.00	3483
Employees, n	1629.46 (2044.45)	1027.00	7.00	24,673
Percentage of Medicare/Medicaid, %	0.45 (0.14)	0.44	0	1
Obesity rate, n	29.44 (3.42)	29.92	20.20	38

The number of hospitals reporting CPT 63 claims increased by 91 from 2012 to 2017. Charges increased from US $2.115 million to US $4.75 million, whereas payments increased from US $1.233 million to US $2.467 million. The proportion of charges paid fluctuated between 45% and 58%. The number of claims increased from 320K to 504K, a 60% increase (Table 3).

Variation across states from 2012 to 2017 for CPT 63 is impressive. The maximum average payment per claim was in Delaware (US $5190.62); however, the number of actual claims was small (5569). New York had the second highest payment per claim (US $5043.79) with 72,186 claims. Texas had the largest number of claims (260,208), yet the average payment per claim was only US $4223.22. On average, 60% of charges were paid (Table 4).

Obesity rates by state have increased from 2012 to 2017 (Table 5). In 2012, the mean obesity rate per state was 27.95%. By 2017, this rate was 30.59%; however, this increase is not weighted by population size. As discussed previously, the aggregate increase for the United States from 2012 to 2017 was 5% (34.9% to 39.6%) [7]. The state data include the District of Columbia (51 observations per year) but are not population weighted.

Table 3. Average statistics by year show the growth in both claims and payments.

Year	Hospitals, n	Total payments, in millions of US $	Total charges, in millions of US $	Total claims, n	Payments/claim, in US $	Charges/claim, in US $
2012	2248	1232.61	2114.56	320,371	3847	6600
2013	2293	1460.80	2698.34	372,155	3925	7251
2014	2306	1517.88	3396.62	410,317	3699	8278
2015	2290	1747.68	3688.72	428,813	4076	8602
2016	2336	2147.48	4247.37	472,004	4550	8999
2017	2339	2466.92	4750.68	504,626	4889	9414

Table 4. Payments, charges, number of claims, payment per claim, charge per claim, and percentage of charges paid by state.

State	Payments, in millions of US $	Charges, in millions of US $	Number of claims, n	Payment/claim, in US $	Charge/claim, in US $	Percentage paid, %
Alaska	21.17	37.94	4387	4826.27	8649.31	55.80
Alabama	232.76	419.16	75,486	3083.47	5552.77	55.50
Arkansas	124.27	200.21	33,087	3755.85	6051.07	62.10
Arizona	175.60	526.67	41,274	4254.39	12,760.42	33.30
California	563.57	1748.02	113,410	4969.31	15,413.29	32.20
Colorado	147.28	440.24	41,537	3545.72	10,598.65	33.50
Connecticut	72.96	139.43	17,179	4247.17	8116.50	52.30
District of Columbia	26.31	53.90	5655	4651.78	9531.16	48.80
Delaware	28.91	25.94	5569	5190.62	4657.34	111.50
Florida	546.44	1543.23	131,442	4157.29	11,740.79	35.40
Georgia	318.64	654.84	79,401	4013.02	8247.29	48.70
Hawaii	12.02	22.03	2802	4291.08	7861.44	54.60
Iowa	131.47	277.09	34,419	3819.62	8050.58	47.40
Idaho	93.02	163.59	23,514	3956.12	6957.22	56.90
Illinois	361.17	706.39	81,960	4406.60	8618.73	51.10
Indiana	361.28	800.98	87,204	4142.93	9185.09	45.10
Kansas	193.40	310.85	39,992	4836.07	7772.78	62.20
Kentucky	215.53	306.93	50,147	4298.00	6120.69	70.20
Louisiana	204.21	423.83	55,726	3664.61	7605.54	48.20
Massachusetts	210.91	267.24	42,609	4950.01	6272.03	78.90
Maryland	83.04	96.72	34,732	2390.89	2784.66	85.90
Maine	69.88	78.25	21,468	3255.23	3645.18	89.30
Michigan	311.32	389.73	70,173	4436.41	5553.81	79.90
Minnesota	191.90	279.98	47,410	4047.73	5905.41	68.50
Missouri	362.53	487.40	83,085	4363.33	5866.30	74.40
Mississippi	159.04	400.04	39,059	4071.72	10,241.97	39.80
Montana	54.19	86.78	14,106	3841.66	6151.89	62.40
North Carolina	503.97	886.27	126,344	3988.87	7014.74	56.90
North Dakota	36.87	43.36	9081	4060.13	4774.58	85.00
Nebraska	94.49	184.44	24,310	3,886.98	7586.98	51.20
New Hampshire	74.81	120.38	19,091	3918.49	6305.40	62.10
New Jersey	169.55	328.00	35,479	4778.93	9244.90	51.70
New Mexico	31.17	68.52	8583	3631.59	7983.19	45.50
Nevada	97.33	190.03	24,617	3953.67	7719.64	51.20
New York	364.09	395.74	72,186	5043.79	5482.21	92.00
Ohio	396.47	743.30	87,427	4534.88	8501.92	53.30
Oklahoma	239.62	494.62	51,639	4640.38	9578.46	48.40
Oregon	180.39	280.10	38,328	4706.54	7307.94	64.40
Pennsylvania	383.17	736.78	85,334	4490.29	8634.05	52.00
Rhode Island	15.94	16.37	3772	4225.09	4339.10	97.40
South Carolina	293.49	492.03	60,977	4813.14	8069.03	59.60

State	Payments, in millions of US $	Charges, in millions of US $	Number of claims, n	Payment/claim, in US $	Charge/claim, in US $	Percentage paid, %
South Dakota	70.42	204.78	17,870	3940.84	11,459.54	34.40
Tennessee	289.03	619.13	79,948	3615.26	7744.18	46.70
Texas	1,098.91	2,450.47	260,208	4223.22	9417.34	44.80
Utah	143.63	138.51	31,168	4608.37	4443.86	103.70
Virginia	231.88	466.58	56,436	4108.69	8267.36	49.70
Vermont	12.44	24.56	3578	3476.00	6864.49	50.60
Washington	264.29	669.87	66,457	3976.84	10,079.74	39.50
Wisconsin	161.13	283.08	39,182	4112.30	7224.71	56.90
West Virginia	117.51	131.28	23,852	4926.48	5504.07	89.50
Wyoming	29.94	40.79	6149	4868.89	6633.27	73.40

Table 5. State statistics for the proportion of the population identified as obese by the Behavioral Risk Factor Surveillance System by year.

Statistic	Year					
	2012	2013	2014	2015	2016	2017
Mean (SD) proportions	27.95 (3.38)	28.65 (3.44)	29.23 (3.42)	29.28 (3.87)	29.78 (3.74)	30.59 (3.86)
Median proportions	27.60	29.40	29.60	29.83	29.92	31.30
Range	14.2	13.8	14.6	16.0	15.39	15.42
Minimum	20.5	21.3	21.3	20.2	22.27	22.64
Maximum	34.7	35.1	35.9	36.2	37.66	38.06
Count	51	51	51	51	51	51

Descriptive Statistics: Categorical Data

Of the 13,769 hospital observations, 3153 were rural and the remaining 10,616 were urban. Most hospital observations were classified as voluntary nonprofits (8866, 64%), whereas proprietary corporations and government entities constituted 3426 (25%) and 1466 (11%), respectively, (11 hospital observations had no ownership specification). Most of the hospital observations (8311, 60%) had no affiliation with medical schools. The vast majority of the observations were from short-term acute care hospitals (13,040, 95%) with nearly all of the remainder (678, 5%) associated with critical access hospitals.

Descriptive Statistics: Correlational Analysis

Hierarchical clustered correlational analysis revealed strong relationships among many of the quantitative variables. Payments and claims are (as to be expected) highly correlated (r=0.9). Most financial and workload metrics are highly correlated as well (eg, net patient revenue and the number of employees; r=0.95). Owing to the large sample size, nearly all correlations are statistically significant at the alpha=.05 level (see Figure 3). The matrix was produced using ggcorrplot [23].

The inclusion of obesity in this study is because of a correlational finding that the number of CPT 63 procedures appears to be influenced by state obesity rates at the aggregate level (Figure 4 [24]). The question, though, is whether this apparent correlation in the logs is sustained when other financial, geographic, technical, and temporal variables are considered.

Figure 3. The correlation matrix depicts the variable relationships. The X's indicate no statistically significant correlation. Owing to the large sample size, nearly all correlations are statistically significant at the alpha=.05 level. The matrix was produced using ggcorrplot.

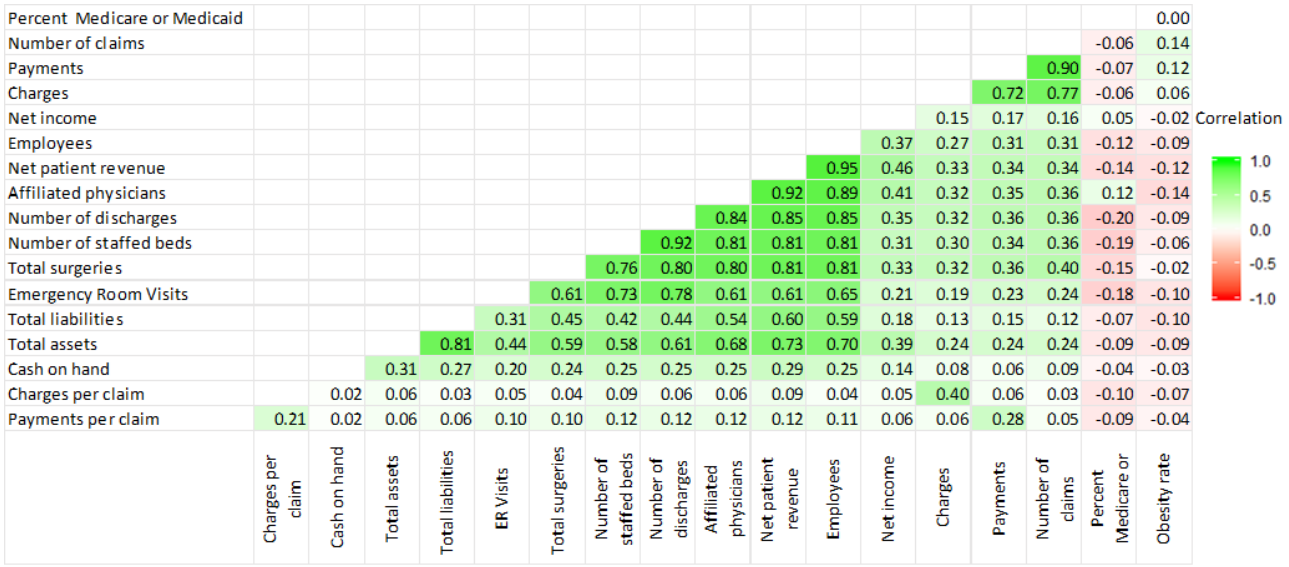
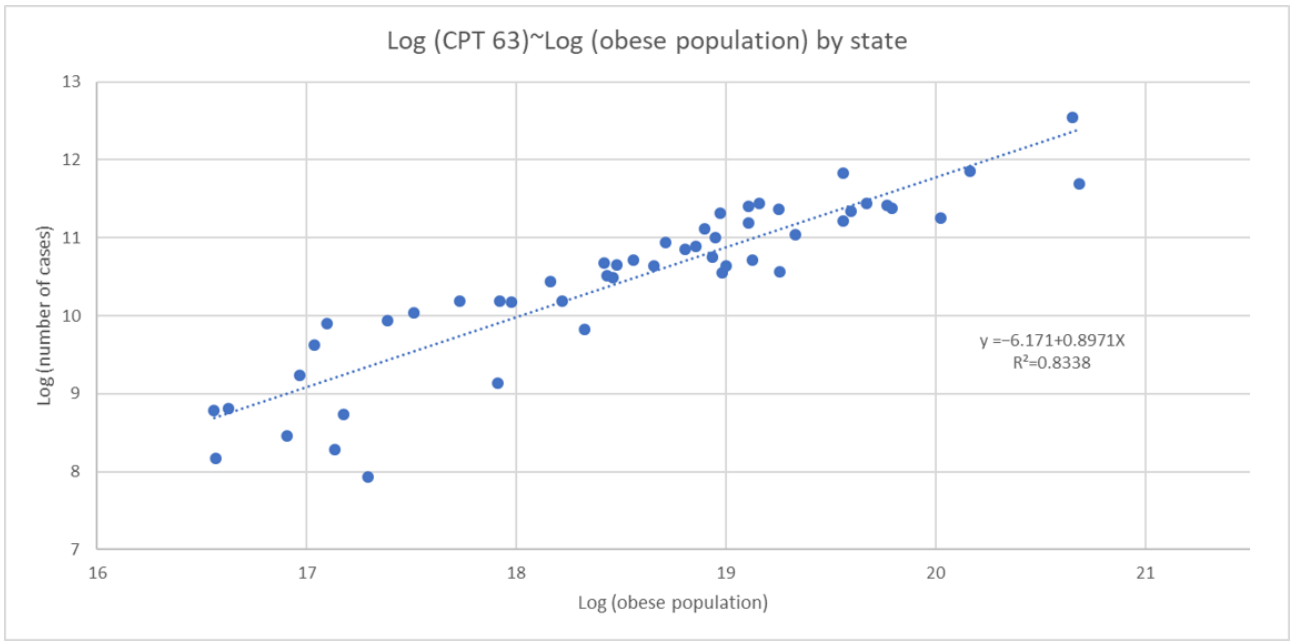


Figure 4. Correlation between the natural logarithm of obesity rates and the natural logarithm of the number of CPT 63 surgeries performed by hospitals.



Exploratory Data Analysis: Feature Engineering and Transformations

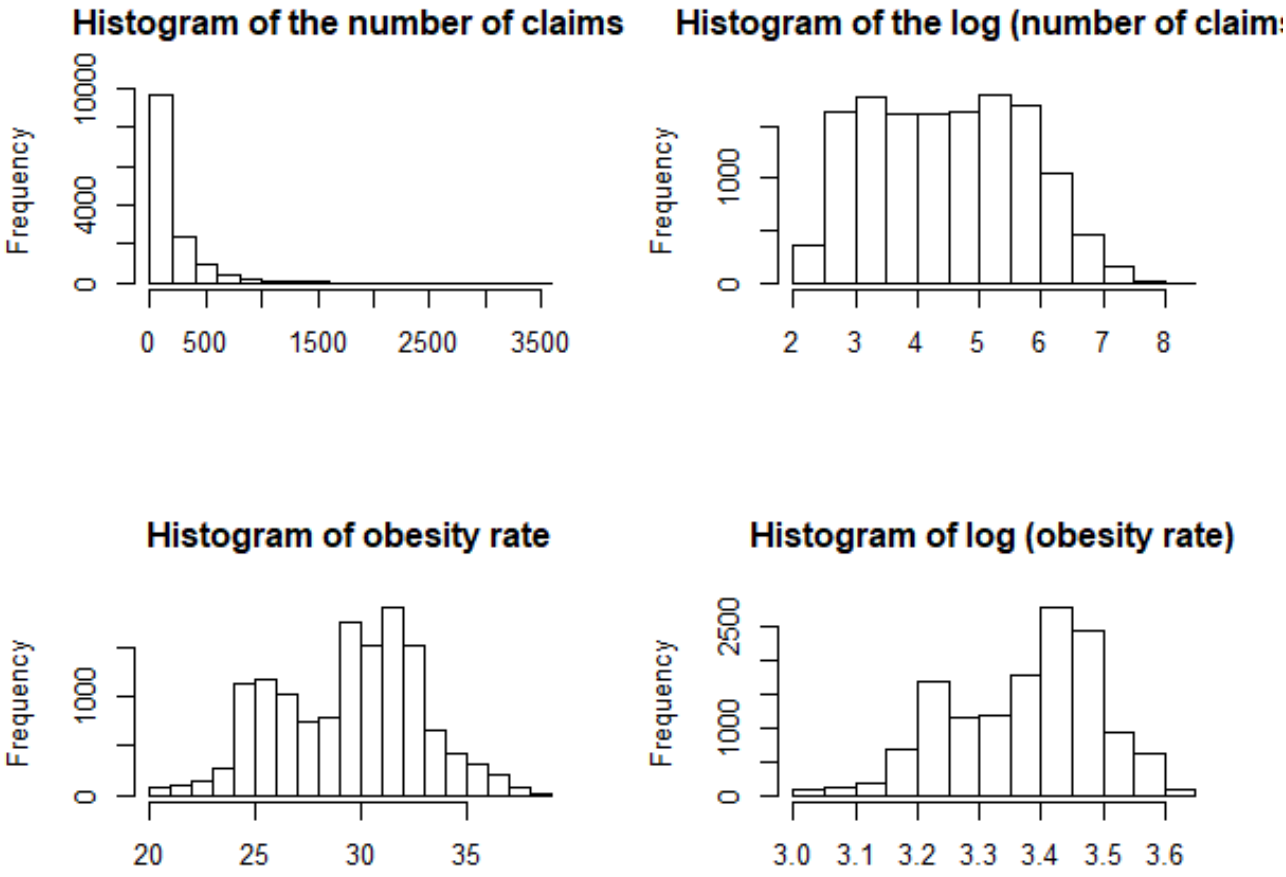
Although random forest regression models are scale invariant, traditional regression techniques such as stepwise, lasso, ridge, and elastic net are not [21]. Investigating transformations to achieve multivariate normality (assuming random-effects regression) is important to meet model assumptions. Furthermore, time series forecasting often benefits from these same techniques [16]. In addition, investigating additional features that might be generated from the existing ones through linear combinations and other methods often results in disentangling collinear variables and finding interesting relationships that might otherwise remain undiscovered.

A multivariate Box-Cox transformation was run using the *car* package in R [25] for all modeled quantitative variables simultaneously after these variables had been location adjusted to make the variables strictly positive, definite, and scale adjusted by dividing by the standard deviation. Multivariate Box-Cox seeks to find power transformations (values of λ for each variable) that make the data multivariate normal enough for use in traditional linear models [26]. These transformations help alleviate the problem of collinearity and address multivariate normal assumptions of random-effects regression. The null hypothesis is that the proposed transformation generated through the transformation is a good fit. The alternative is that it is not a good fit. The proposed transformation was a vector of primarily natural logarithms (values near zero) with some exceptions. The likelihood ratio test resulted in a *P* value >.99, indicating that the assumption of multivariate normality cannot be rejected. The actual vector of transformations follows: λ={0.1, 0.3, 0.33, 0.38, 0.21, −0.07, 1.66, 0.76, −0.03, 0.55, 0.28, 0.17, 1.03, 1.04} for x={number

of claims, number of staffed beds, number of discharges, ER visits, total surgeries, net patient revenue, net income, cash, total assets, total liabilities, affiliated physicians, employees, percent Medicare/Medicaid, obesity rate}, respectively.

Univariate histograms for the number of claims and obesity pre- and posttransformation are in Figure 5. The transformed graph of the number of claims shows some slight skew but is otherwise unremarkable. However, the graph of obesity rates is telling. Although the transformation fails to reject the assumption of multivariate normality, the obesity graph is bimodal. It is possible that linear-in-parameter models will not be able to correctly fit the importance of this variable, whereas tree-based models will find patterns.

Figure 5. The untransformed and transformed histograms of the number of claims and the obesity rate variables.



Geospatial Analysis Results: Zip Code Unit of Analysis

Geospatial heat map analysis of CPT 63 number of claims by year and parsed by zip code is shown in panels (Figure 6). The maximum scale is approximately 7000 claims for each diagram to allow for comparison across years. Multimedia Appendix 1 shows this analysis in video format.

In 2012, there was very little high-intensity activity (Houston and Dallas, Texas, primarily, with some activity in the Carolinas). The Eastern seaboard has activity, but it is not intense, and the Western seaboard has minimal activity, except near Seattle.

By 2013, the Eastern seaboard (particularly New Jersey) has increased in intensity, and the areas around Chicago and Salem, Oregon, are emerging as well. Houston and Dallas remain the most intense regions for the number of claims.

In 2014, the number of claims in Seattle and San Antonio, Texas, shifted these cities to high-intensity areas (some red visible). It must be noted that 3 of the 4 cities with visible red tint are in Texas (San Antonio, Dallas, and Houston).

The year 2015 saw increasing intensity in both the New Jersey area and Chicago, Illinois. These 2 areas joined Houston, Dallas, San Antonio, and Seattle as high-intensity claims areas. Despite their populations, neither California nor Florida experienced the claims intensity of Texas.

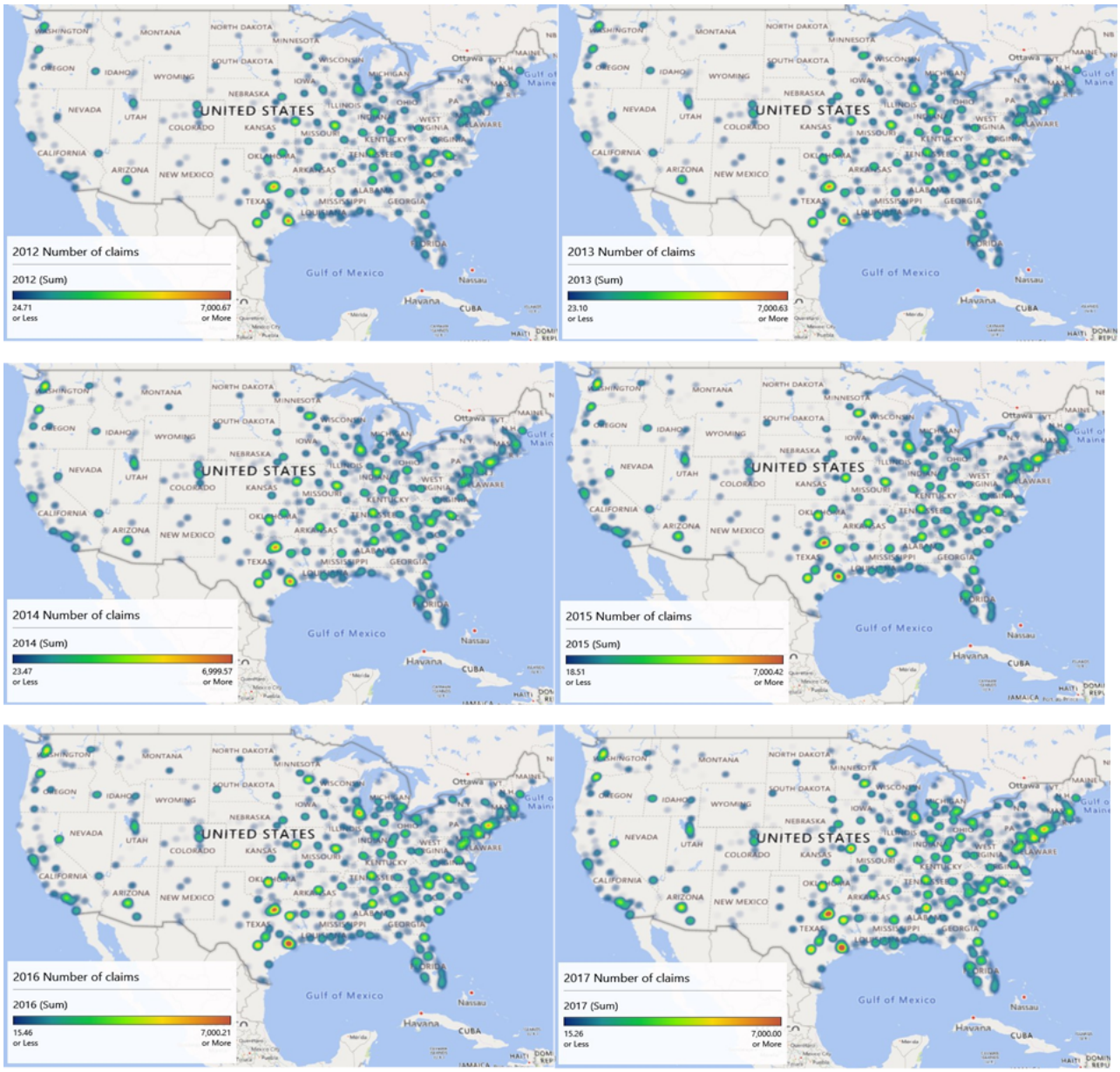
Houston, Dallas, San Antonio, Seattle, Longview (Texas), Oklahoma City, the New Jersey area, and St. Louis were the notable areas of high intensity in 2016. The California coast became more intense along with Salt Lake City.

By 2017, the Eastern seaboard intensified (New Jersey, Delaware, Pennsylvania, Virginia, and District of Columbia)

along with Phoenix, Arizona, and Atlanta, Georgia areas. The most intense areas remained Houston and Dallas.

Overall, the maps may suggest small area variations in practice patterns [27]. Although California and Florida have large populations, none of their major population centers reached the high-intensity scale of major cities in Texas. Furthermore, the Eastern seaboard's increasing intensity suggests that something has changed. The questions then become are these changes in demand forecastable and how might they be explained.

Figure 6. Geospatial analysis of all CPT 63 claims from 2012 through 2017.



Forecasting Results

Number of Claims

Being able to forecast demand is necessary for decision makers to investigate both supply- and demand-side interventions. To that end, HTS for state, census bureau region, and the nation using both ETS and ARIMA components were built on 2012-2015 training dataset and compared with the 2016-2017 test set using the *hts* package in R [16]. Bottom-up, top-down, middle-out, and combination approaches to this forecasting were analyzed.

The ETS models performed better on the test set in terms of both variance and bias as shown (Table 6), and the middle-out model performed better on all bias (mean error and mean percentage error) as well as variance (root mean squared error, mean absolute error, and mean absolute percentage error) metrics. The overall forecast from the ETS middle-out model for the unobserved years {2016, 2017} was {454,720.3, 482,049.9}, whereas the actual overall claims were {464,323, 497,325}, resulting in mean absolute percent error (MAPE) of {2.0%, 3.1%}. Table 7 illustrates the forecast and actual number of claims at the region-level hierarchy for the best performing model, whereas Table 8 provides the state-by-state forecasts.

Table 6. The performance metrics of the various hierarchical models show that the exponential, trend, seasonality middle-out model performed best on the test set.

Model	Mean error	Root mean squared error	Mean absolute error	Mean percent error (%)	Mean absolute percent error (%)
ETS[a]-bottom up	860.42	2905.64	1371.81	3.16	10.71
ETS-top down	688.26	2423.58	1266.13	3.03	10.54
ETS-middle out	611.75	2256.70	1219.59	2.70	10.50
ETS-combination	682.27	2404.41	1235.04	2.25	9.56
ARIMA[b]-bottom up	5732.63	16496.72	5762.20	24.93	26.65
ARIMA-top down	5214.61	14953.59	5312.37	23.26	25.66
ARIMA-middle out	4606.38	13420.78	4799.88	20.70	24.20
ARIMA-combination	5159.04	14782.37	5259.67	20.65	25.92

[a]ETS: exponential, trend, seasonality.

[b]ARIMA: autoregressive integrated moving average.

Table 7. Region-level forecasts demonstrate small error. The average mean absolute percent error (MAPE) for {2016, 2017} was {3.4%, 6.2%}, respectively.

Region	2016 forecast	2016 actual	2017 forecast	2017 actual	MAPE 2016 (%)	MAPE 2017 (%)
East North Central	67,019	67,672	70,630	72,720	1.0	2.9
East South Central	42,765	42,747	44,477	47,107	0.0	5.6
Middle Atlantic	35,044	37,532	38,309	42,860	6.6	10.6
Mountain	35,840	35,254	38,211	38,341	1.7	0.3
New England	17,763	19,476	17,763	20,733	8.8	14.3
Pacific	45,015	43,647	49,122	43,865	3.1	12.0
South Atlantic	94,650	96,093	100,266	103,546	1.5	3.2
West North Central	45,653	46,550	48,075	48,600	1.9	1.1
West South Central	70,972	75,352	75,197	79,553	5.8	5.5

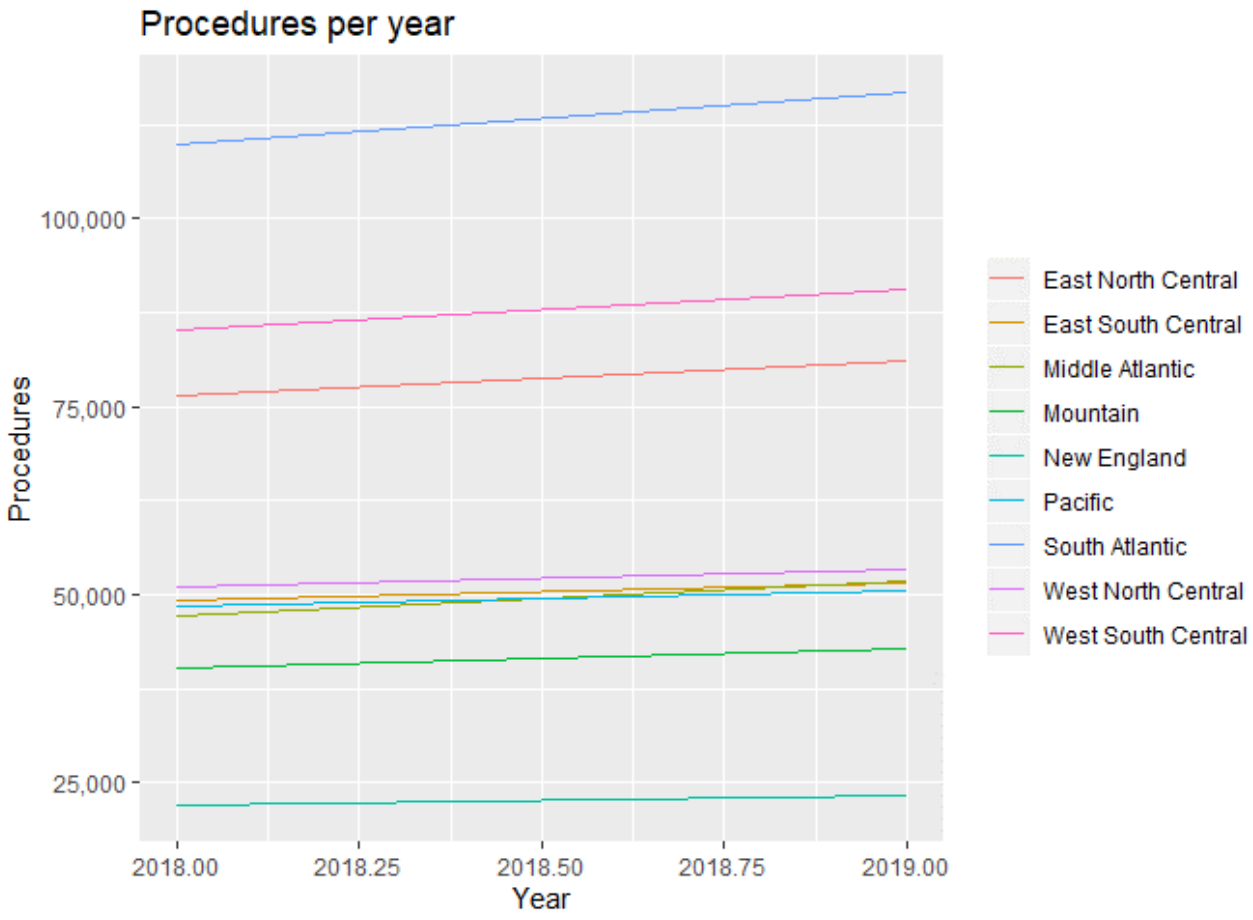
Table 8. Forecasts produced by the exponential, trend, seasonality middle-out model by state for 2016 and 2017 have an average mean absolute percent error (MAPE) of 10.1% and 13.2%, respectively.

State	2016 forecast	2016 actual	2017 forecast	2017 actual	Mean absolute error 2016 (%)	Mean absolute error 2017 (%)
Alaska	773	1,132	825	784	31.7	5.2
Alabama	13,709	14,029	14,658	15,697	2.3	6.6
Arkansas	6550	6193	7077	6150	5.8	15.1
Arizona	8450	7585	9438	9624	11.4	1.9
California	22,060	21,402	24,242	21,859	3.1	10.9
Colorado	7089	6736	7081	6512	5.2	8.7
Connecticut	2539	3103	2446	3289	18.2	25.6
Dist. of Columbia	1052	1015	1070	1041	3.6	2.8
Delaware	702	1129	683	1286	37.8	46.9
Florida	24,991	23,611	26,870	25,417	5.8	5.7
Georgia	15,324	14,293	16,372	14,093	7.2	16.2
Hawaii	369	451	370	527	18.2	29.8
Iowa	5613	5698	5624	6069	1.5	7.3
Idaho	3775	4755	3920	4793	20.6	18.2
Illinois	15,590	15,793	16,474	15,168	1.3	8.6
Indiana	14,839	14,839	14,850	14,648	0.0	1.4
Kansas	6943	7792	7382	7969	10.9	7.4
Kentucky	8061	7805	8044	9239	3.3	12.9
Louisiana	10,682	9817	11,519	11,044	8.8	4.3
Massachusetts	7251	8424	7533	9556	13.9	21.2
Maryland	6196	6458	6475	6785	4.1	4.6
Maine	3532	3707	3402	3345	4.7	1.7
Michigan	13,843	12,998	15,005	14,914	6.5	0.6
Minnesota	10,768	9,499	12,317	9507	13.4	29.6
Missouri	13,157	14,155	13,181	15,378	7.1	14.3
Mississippi	6943	7562	7252	7301	8.2	0.7
Montana	2954	2638	3116	2108	12.0	47.8
North Carolina	21,173	22,429	21,527	24,314	5.6	11.5
North Dakota	1839	1672	2017	2013	10.0	0.2
Nebraska	4437	4036	4576	4210	9.9	8.7
New Hampshire	3000	3073	2890	3382	2.4	14.5
New Jersey	6643	7067	7102	6659	6.0	6.7
New Mexico	1351	1570	1350	1755	13.9	23.1
Nevada	5067	4497	5606	5320	12.7	5.4
New York	12,876	14,206	14,014	15,442	9.4	9.2
Ohio	15,944	16,977	17,075	19,643	6.1	13.1
Oklahoma	9613	10,214	10,421	10,386	5.9	0.3
Oregon	7875	7,111	8454	7574	10.7	11.6
Pennsylvania	15,524	16,259	17,193	20,759	4.5	17.2
Rhode Island	841	695	914	750	21.0	21.9
South Carolina	11,050	10,921	11,858	13,494	1.2	12.1

State	2016 forecast	2016 actual	2017 forecast	2017 actual	Mean absolute error 2016 (%)	Mean absolute error 2017 (%)
South Dakota	2896	3698	2977	3454	21.7	13.8
Tennessee	14,052	13,351	14,523	14,870	5.3	2.3
Texas	44,127	49,128	46,181	51,973	10.2	11.1
Utah	6220	6077	6765	6811	2.4	0.7
Virginia	10,543	12,143	11,731	12,020	13.2	2.4
Vermont	600	474	578	411	26.6	40.6
Washington	13,937	13,551	15,231	13,121	2.8	16.1
Wisconsin	6803	7065	7227	8347	3.7	13.
West Virginia	3619	4094	3680	5096	11.6	27.8
Wyoming	935	1396	934	1418	33.0	34.1

HTS with the middle-out approach and ETS methods was refit on the entire dataset to generate forecasts. Figure 7 shows the regional forecasts for 2018 and 2019. The East North Central region of the country is likely to experience the largest growth in claims. The overall demand for 2018 and 2019 is forecasted to be {529,777, 562,023}, which represents growth of 6.52% growth in the first year (from 497,325 procedures in 2017 to 529,777 in 2018) and 13.00% by 2019 (from 497,325 procedures in 2017 to 562,023 in 2019). At US $5000 average per claim (a simple linear model would suggest US $4910 in 2018 and US $5123 in 2019), the net increase in cost for 2018 would be US $162.2 million for 2018 and US $323.9 million for 2019. The next question becomes what explains the predicted growth of these claims other than possibly practice variation.

Figure 7. Regional forecasts generated by the hierarchical time series middle-out model with exponential, trend, seasonality components.



Explanatory Modeling Results

To investigate explanatory variables, several models were explored. Stepwise regression for the number of claims at the hospital level using the transformed variables and an 80% training set was successfully able to predict the number of claims on the withheld test set with some accuracy (adjusted R^2=0.39

on the training set and adjusted R^2=0.38 on the test set). This indicates that the sum of squared regression accounted for 38% of the variance of the sum of squared total on the test set. Payments and charges were excluded from the analysis as they are necessarily functions of claims. The variables evaluated were the number of staffed beds, discharges, surgeries, net patient revenue, net income, total assets, total liabilities, affiliated physicians, employees, percentage Medicare/Medicaid, state, year, urban/rural status, ownership, medical school status, and hospital type. Table 9 provides the remaining variables generated from the stepwise regression at the hospital unit of analysis. It should be noted that obesity did not remain in the final model.

Stepwise regression for the number of claims with data aggregated (mean) by state and by year (N=306 observations, 51 states/territories × 6 years) resulted in an impressive model using an 80% training set to predict a 20% withhold set. The adjusted R^2 was 0.87 on the training set and 0.77 on the test set after dropping insignificant variables from the analysis. The variables in this model included state, year, number of discharges, and total liabilities (a parsimonious model; Table 10). Again, there is no evidence that obesity rates are predictive of CPT 63 surgery in this model.

Lasso, ridge, and elastic net regression models were able to predict the unaggregated test set with some accuracy (R^2=0.38, 0.37, 0.38, respectively.) None of these penalty-weighted models improved upon the stepwise analysis significantly, although elastic net tied. Obesity was not retained in these models. For the aggregated set (state and year), the associated R^2 were 0.78, 0.75, and 0.78, respectively. The lasso and elastic net models were slightly superior to the stepwise regression model (Figure 5). The top 10 variables by effect size in the state-aggregated elastic net model are shown in Table 11. The effect size of obesity was near zero (0.0098). If one were to make a conclusion using traditional and constrained linear models, obesity would not be a factor for explaining the number of claims; however, random forests would prove otherwise.

Table 9. Variables below from the stepwise regression predicted a withhold set with adjusted R^2=0.38.

Variable	Sum of squares	Mean squared error	*F* value (*df*)	*P* value
Staffed beds	95.78	95.78	2423.29 (1)	<.001
Discharges	21.93	21.93	554.77 (1)	<.001
ER visits	17.50	17.50	442.68 (1)	<.001
Surgeries	41.15	41.15	1041.03 (1)	<.001
Net patient revenue	5.57	5.57	140.99 (1)	<.001
Net income	2.53	2.53	63.97 (1)	<.001
Total liabilities	4.45	4.45	112.70 (1)	<.001
Affiliated physicians	0.24	0.24	6.08 (1)	<.01
Employees	6.28	6.28	158.78 (1)	<.001
Percentage Medicare/Medicaid	1.16	1.16	29.34 (1)	<.001
State	52.35	1.05	26.49 (50)	<.001
Year	8.09	1.62	40.91 (5)	<.001
Urban rural status	1.99	1.99	50.43 (1)	<.001
Ownership	10.76	0.90	22.68 (12)	<.001
Medical school affiliation	2.16	0.54	13.63 (4)	<.001
Hospital type	0.64	0.13	3.24 (5)	<.01

Table 10. Variables in the analysis by state and by year.

Variable	Sum of squares	Mean squared error	*F* value (*df*)	*P* value
State	1.16	0.02	26.40 (50)	<.01
Year	0.17	0.03	39.16 (5)	<.01
Discharges	0.06	0.06	63.19 (1)	<.01
Net income	0.004	0.004	4.45 (1)	.04
Total liabilities	0.004	0.004	4.65 (1)	.03

Table 11. Top 10 coefficients by effect size of the elastic net.

Variable	Coefficient
Total assets	−1.539
Net patient revenue	−1.044
Number of staffed beds	0.212
Number of discharges	0.186
New Jersey	−0.178
Total surgeries	0.162
New York	−0.159
California	−0.145
Delaware	−0.143
Employees	−0.136

Gradient-boosted random forests with hyperparameter tuning outperformed all models: stepwise, lasso, ridge, elastic net regression. On the unaggregated withhold set, a well-pruned model (depth 4) with 2000 epoch runs and a slow learning rate of 0.1 accounted for more than 78.5% of the variability (R^2=0.79) on the unobserved test set. Comparing this value with the approximately 38% variability accounted for in the other models suggests that the random forest model is superior. Figure 8 is a plot of the gain (the average improvement when the feature is used in a tree) for the top 5 items in the importance matrix, whereas Figure 9 is a plot of the cover (the average proportion of samples affected by splitting using this feature) for the top 5 items of the unaggregated model. These figures illustrate that obesity is one of the prominent features in both gain and cover of the unaggregated model.

Figure 8. Gain plot for the top 5 variables, unaggregated model.

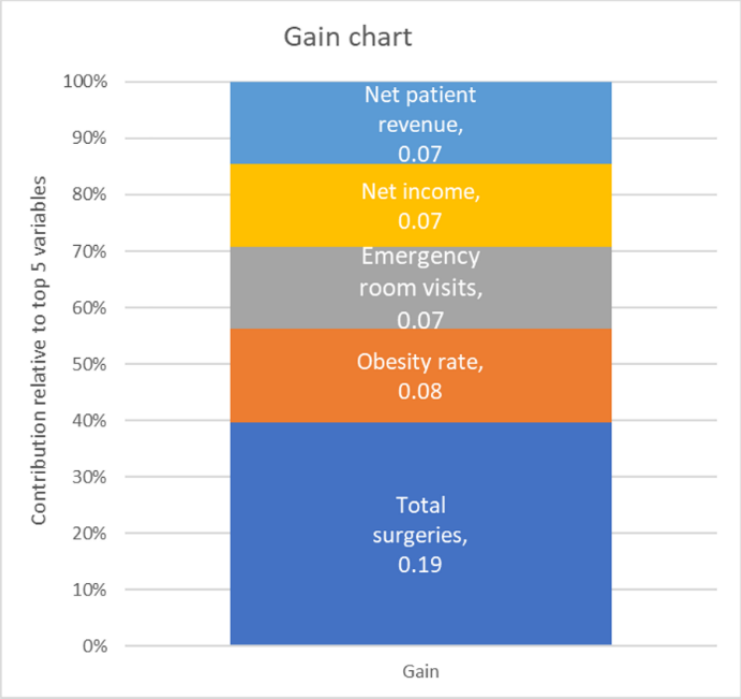
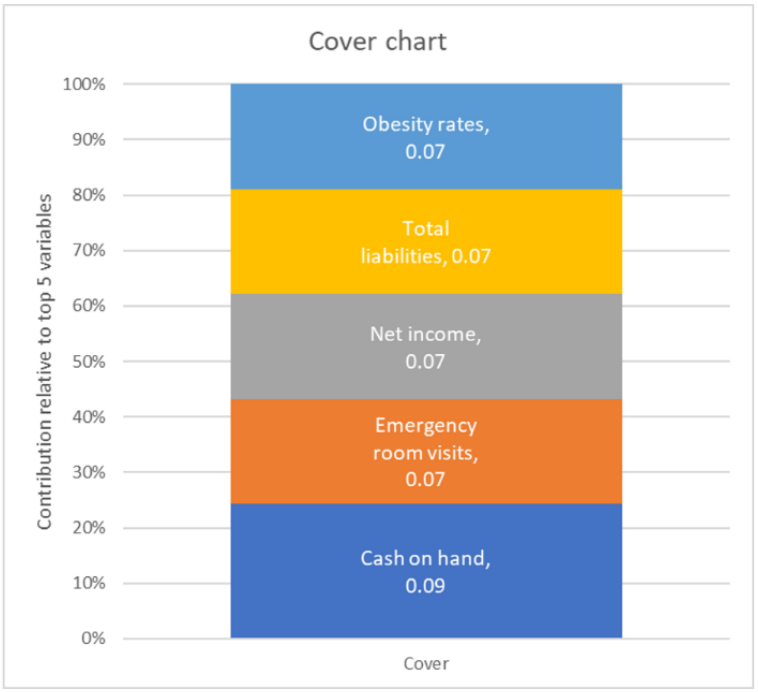


Figure 9. Cover plot for the top 5 variables, unaggregated model.



Despite the exceptional gains of the extreme gradient-boosted random forests on the unaggregated, hospital-level data, the application of hyperparameter-tuned models to the aggregated data (by state and year) yielded only nominal improvement over the constrained regression methods, possibly because of the smaller sample due to aggregation. A well-pruned model (depth=3) after 3000 epochs with a slow learning rate (0.1) achieved an R^2 of 0.80. The gain and cover graphs are shown in Figures 10 and 11, and obesity rate is the most important feature at the state-aggregated level.

Most importantly, the gradient-boosted random forests identified obesity as the second most important factor for gain at the hospital level and as the most important factor for both gain and cover at the state level of analysis. Furthermore, the gradient-boosted random forests performed better than any other model considered on a blinded test set.

Figure 10. Gain plot for the top 5 variables, aggregated model.

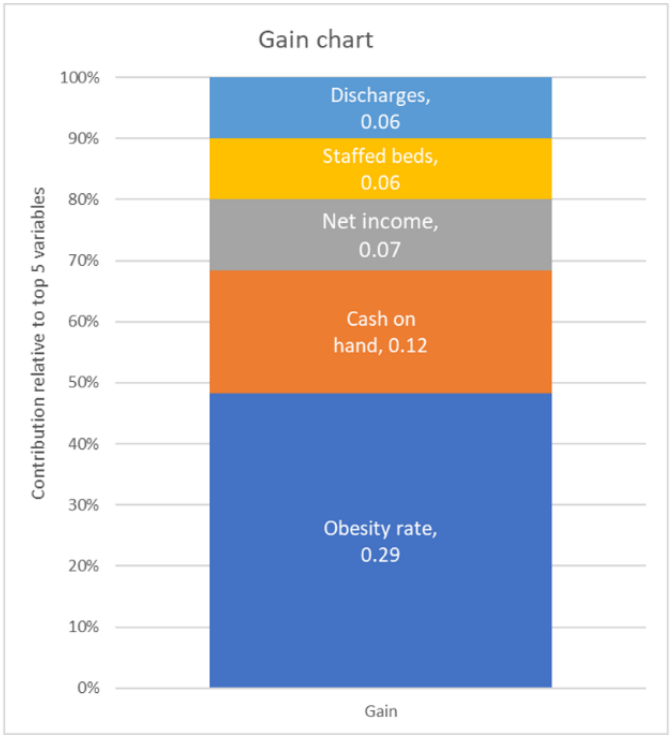
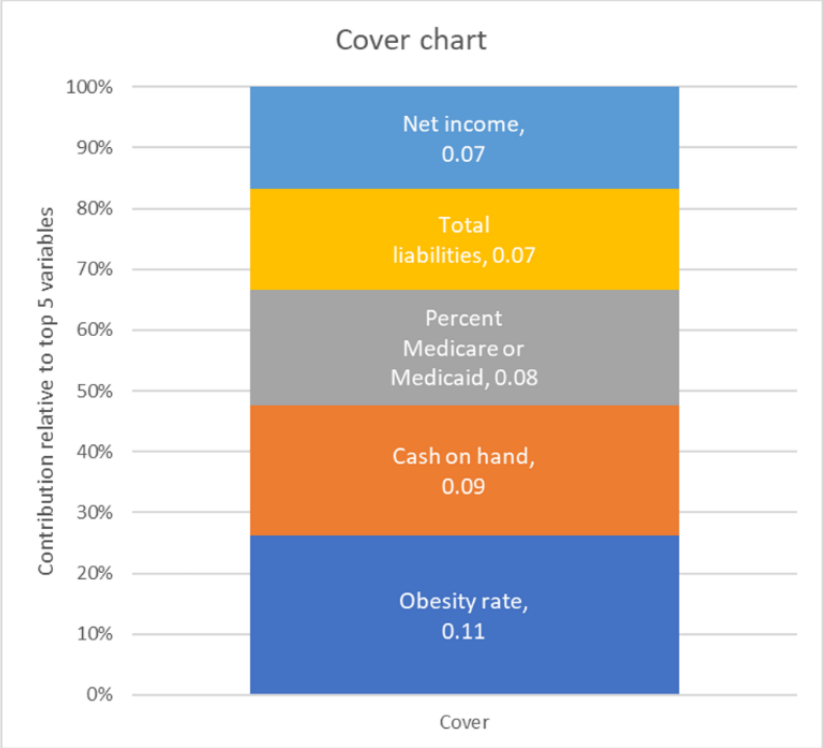


Figure 11. Cover plot for the top 5 variables, aggregated model.



Discussion

Principal Findings

In this analysis, we evaluated the location, magnitude, and reasons for the growth of CPT 63 back surgeries in the United States. The GIS heat map analysis shows large-scale growth, particularly in the Northeastern region of the United States, and sustained activity in Texas. The entirety of the Eastern seaboard has seen growth in these procedures, and the associated increased cost is estimated to be US $323.9 million by the end of 2019.

The principal findings of this study are described here. Each of the following results includes a discussion of significance and (if appropriate) policy:

1. The Northeastern seaboard is likely to see continued growth in CPT 63 procedures. The implication for states in this region is that they may see more unplanned expenditures on health care, affecting their budgets. Furthermore, cost controls and reduction of practice variation based on evidence will become more important.
2. The cost associated with these procedures is outstripping inflation and will likely result in national expenditures in the triple-digit billions. The federal government may need to evaluate its own evidence-based, best practice policies associated with funding of procedures that link selected interventions with outcomes and that reasonably limit reimbursement.
3. Interstate practice variation appears to be extreme. For example, large population centers in California have fewer claims than large population centers in Texas. States should also investigate intrastate variation.
4. Hierarchical forecasts suggest an increase in the number of claims of 6.5% for 2018 and 13% in 2019. The initial models were built on a blind test set and performed well. These types of forecasts are reasonably effective for claims analysis.
5. Explanatory regression models for nation-level claims data had only some success in internal predictions. These models excluded obesity as a predictor. Regression models were more successful at predicting aggregated state/year models, though. These traditional models should be abandoned in favor of random forests.
6. Extreme gradient-boosted random forest models were highly successful in predicting both hospital-level unit of analysis number of claims and aggregate-level claims on an unobserved test set. These models identified obesity as an important factor in estimating the number of claims. Furthermore, the use of these models underscores that even after multivariate transformations, nonlinear functions may exist in modeled data. Random forests unearthed patterns not visible to regression and constrained regression models.

Limitations

There are many limitations in this work. First, the algorithms used by Definitive Healthcare to extrapolate CMS data to *all-payor* data are not divulged. This omission is problematic for verification but understandable because of parochial concerns. Second, only ETS and ARIMA models were considered for the HTS fitting as these models are implemented in the R HTS package. There are an infinite number of models for forecasting, including random forest time series that might have performed better. Third, the explanatory variables are limited to those tracked by CMS and the BRFSS.

Conclusions

Hospital-based back surgeries are likely to increase dramatically over the next several years, yet the supply of neurosurgeons is constant. With that increase, the cost of the procedures (mostly borne by third-party payers) will increase as well. Practice variation appears to be prevalent across the country; however, obesity itself is a factor that must be considered as a significant influence. Policy interventions must be considered at many levels.

Clinical practice variation is something that may require intervention at the federal level. For example, a study in Scandinavia found significant differences among Norway, Sweden, and Denmark in the use of concomitant arthrodesis without any difference in treatment efficacy, increasing the cost without improving outcomes [28]. Controlling costs across states may require federal (and state) reimbursement interventions and incentives.

States should continue educational and financial interventions targeting obesity in adults and children. As the obesity epidemic continues to grow, the medical intervention costs are likely to grow accordingly. Furthermore, states should evaluate county-to-county practice variation as these variations often increase cost without improving quality [27].

Local interventions should consider the targeting of food deserts (urban areas where fresh, quality food is difficult to find) for eradication as well as educational interventions. Several studies have shown that the food environment is directly linked to obesity [29-31]. Eliminating or at least reducing the number of food deserts requires incentivizing grocery stores to populate areas where it may not be as lucrative because of poverty or demand.

Insurance companies themselves have a vested interest in both reducing obesity and controlling practice variation. Obesity is linked to numerous health disorders such as heart disease, type 2 diabetes, and bone and joint disease [32], any of which may result in additional costs to the health care system and insurer. Funding prevention efforts and establishing policies to reduce practice area variation are likely to benefit them as well as the population health over time.

Federal, state, and local policy makers need to address the increasing obesity epidemic and the likely associated increase in demand for back surgeries. The implications of not doing so are increased cost, questionable quality/cost trade-offs, and reduced access because of the small and steady number of available neurosurgeons. The *fattening of America* and the costs associated with it are likely to continue increasing otherwise.

Conflicts of Interest

None declared.

Multimedia Appendix 1
This video depicts heat maps for the number of claims from 2012 through 2017.
[MP4 File (MP4 Video), 4334 KB - jmir_v21i10e14609_app1.mp4]

References

1. American Association of Neurological Surgeons. 2012. Ensuring an Adequate Neurosurgical Workforce for the 21st Century URL: https://www.aans.org/pdf/Legislative/Neurosurgery%20IOM%20GME%20Paper%2012%2019%2012.pdf [accessed 2019-09-03] [WebCite Cache ID 788BVM51T]
2. Greenwood B. Houston Chronicle. 2018. What Is the Job Outlook for a Neurosurgeon? URL: https://work.chron.com/job-outlook-neurosurgeon-20074.html [accessed 2019-05-05] [WebCite Cache ID 788BjABT7]
3. Definitive Healthcare: Healthcare Analytics & Provider Data. 2019. URL: https://www.definitivehc.com/ [accessed 2019-05-05] [WebCite Cache ID 788Bun3Wb]
4. SuperCoder: Medical Coding & Billing Tools. 2019. CPT® Codes — Surgical Procedures on the Spine and Spinal Cord URL: https://www.supercoder.com/cpt-codes-range/1844 [accessed 2019-08-31] [WebCite Cache ID 788C25UMY]
5. Azad TD, Vail D, O'Connell C, Han SS, Veeravagu A, Ratliff JK. Geographic variation in the surgical management of lumbar spondylolisthesis: characterizing practice patterns and outcomes. Spine J 2018 Dec;18(12):2232-2238. [doi: 10.1016/j.spinee.2018.05.008] [Medline: 29746964]
6. Alvin MD, Lubelski D, Alam R, Williams SK, Obuchowski NA, Steinmetz MP, et al. Spine surgeon treatment variability: the impact on costs. Global Spine J 2018 Aug;8(5):498-506 [FREE Full text] [doi: 10.1177/2192568217739610] [Medline: 30258756]
7. The State of Obesity – Better Policies for a Healthier America. 2019. National Obesity Rates & Trends URL: https://www.stateofobesity.org/obesity-rates-trends-overview/ [accessed 2019-08-31] [WebCite Cache ID 788CDkupM]
8. Finkelstein EA, Trogdon JG, Cohen JW, Dietz W. Annual medical spending attributable to obesity: payer-and service-specific estimates. Health Aff (Millwood) 2009;28(5):w822-w831. [doi: 10.1377/hlthaff.28.5.w822] [Medline: 19635784]
9. Chan AK, Bisson EF, Bydon M, Glassman SD, Foley KT, Potts EA, et al. Obese patients benefit, but do not fare as well as nonobese patients, following lumbar spondylolisthesis surgery: an analysis of the quality outcomes database. Neurosurgery 2018 Dec 12:- (epub ahead of print). [doi: 10.1093/neuros/nyy589] [Medline: 30541141]

10. Zhang TT, Liu Z, Liu YL, Zhao JJ, Liu DW, Tian QB. Obesity as a risk factor for low back pain: a meta-analysis. Clin Spine Surg 2018 Feb;31(1):22-27. [doi: 10.1097/BSD.0000000000000468] [Medline: 27875413]
11. Chronic Disease and Health Promotion Data & Indicators - CDC. 2019. BRFSS: Table of Overweight and Obesity (BMI) URL: https://chronicdata.cdc.gov/Behavioral-Risk-Factors/BRFSS-Table-of-Overweight-and-Obesity-BMI-/fqb7-mgjf [accessed 2019-05-05] [WebCite Cache ID 788CKzB88]
12. Lee DC, Yi SS, Athens JK, Vinson AJ, Wall SP, Ravenell JE. Using geospatial analysis and emergency claims data to improve minority health surveillance. J Racial Ethn Health Disparities 2018 Aug;5(4):712-720 [FREE Full text] [doi: 10.1007/s40615-017-0415-4] [Medline: 28791583]
13. MacQuillan EL, Curtis AB, Baker KM, Paul R, Back YO. Using GIS mapping to target public health interventions: examining birth outcomes across GIS techniques. J Community Health 2017 Aug;42(4):633-638. [doi: 10.1007/s10900-016-0298-z] [Medline: 27885583]
14. Côté MJ, Smith MA. Forecasting the demand for radiology services. Health Syst (Basingstoke) 2018;7(2):79-88 [FREE Full text] [doi: 10.1080/20476965.2017.1390056] [Medline: 31214340]
15. Brookmeyer R, Johnson E, Ziegler-Graham K, Arrighi HM. Forecasting the global burden of Alzheimer's disease. Alzheimers Dement 2007 Jul;3(3):186-191. [doi: 10.1016/j.jalz.2007.04.381] [Medline: 19595937]
16. Hyndman R, Lee A, Wang E, Wickramasuriya S. RDDR. 2018. HTS-package: Hierarchical and Grouped Time Series URL: https://rdrr.io/cran/hts/man/hts-package.html [accessed 2019-08-31]
17. The R Development Core Team. Servidor de software libre de la Universidad de Zaragoza. 2018. R: A Language and Environment for Statistical Computing URL: https://www.r-project.org/ [accessed 2019-09-03]
18. MacNab YC. Hierarchical Bayesian modeling of spatially correlated health service outcome and utilization rates. Biometrics 2003 Jun;59(2):305-316. [doi: 10.1111/1541-0420.00037] [Medline: 12926715]
19. Hyndman RJ, Athanasopoulos G. Forecasting: Principles and Practice. Second Edition. Boston, Massachusetts: OTexts; 2019.
20. Holt CC. Forecasting seasonals and trends by exponentially weighted moving averages. Int J Forecast 2004 Jan;20(1):5-10. [doi: 10.1016/j.ijforecast.2003.09.015]
21. Hastie T, Tibshirani R, Friedman J. TheElements of Statistical Learning: Data Mining, Inference, and Prediction. Second Edition. New York: Springer; 2019.
22. Chen T, He T, Benesty M, Khotilovich V, Tang Y, Cho H. XGBoost Documentation. 2018. XGBoost: Extreme Gradient Boosting URL: https://xgboost.readthedocs.io/en/latest/ [accessed 2019-08-31]
23. Kassambara A. Statistical Tools for High-Throughput Data Analysis. 2018. ggcorrplot: Visualization of a Correlation Matrix Using ggplot2 URL: http://www.sthda.com/english/wiki/ggcorrplot-visualization-of-a-correlation-matrix-using-ggplot2 [accessed 2019-08-31]
24. Softonic. 2016. Microsoft Excel 2016 URL: https://microsoft-excel-2016.en.softonic.com/ [accessed 2019-08-31]
25. Fox J, Weisberg S. An R Companion to Applied Regression. Second Edition. Thousand Oaks, CA: Sage Publications; 2011.
26. Box GE, Cox DR. An analysis of transformations. J R Stat Soc Series B Stat Methodol 2018 Dec 5;26(2):211-243 [FREE Full text] [doi: 10.1111/j.2517-6161.1964.tb00553.x]
27. Wennberg J, Gittelsohn N. Small area variations in health care delivery. Science 1973 Dec 14;182(4117):1102-1108. [doi: 10.1126/science.182.4117.1102] [Medline: 4750608]
28. Lønne G, Fritzell P, Hägg O, Nordvall D, Gerdhem P, Lagerbäck T, et al. Lumbar spinal stenosis: comparison of surgical practice variation and clinical outcome in three national spine registries. Spine J 2019 Jan;19(1):41-49. [doi: 10.1016/j.spinee.2018.05.028] [Medline: 29792994]
29. Morland K, Wing S, Roux AD. The contextual effect of the local food environment on residents' diets: the atherosclerosis risk in communities study. Am J Public Health 2002 Nov;92(11):1761-1767. [doi: 10.2105/ajph.92.11.1761] [Medline: 12406805]
30. Morland K, Roux AV, Wing S. Supermarkets, other food stores, and obesity: the atherosclerosis risk in communities study. Am J Prev Med 2006 Apr;30(4):333-339. [doi: 10.1016/j.amepre.2005.11.003] [Medline: 16530621]
31. Holsten JE. Obesity and the community food environment: a systematic review. Public Health Nutr 2009 Mar;12(3):397-405. [doi: 10.1017/S1368980008002267] [Medline: 18477414]
32. Stanford Health Care (SHC) - Stanford Medical Center. 2019. Effects of Obesity URL: https://stanfordhealthcare.org/medical-conditions/healthy-living/obesity.html [accessed 2019-08-31]

Abbreviations

ARIMA: autoregressive integrated moving average models
BRFSS: Behavioral Risk Factor Surveillance System
CPT: Current Procedural Terminology
CMS: Centers for Medicare and Medicaid Services
ETS: exponential, trend, seasonality

HTS: hierarchical time series
MAPE: mean absolute percent error

Edited by G Eysenbach; submitted 05.05.19; peer-reviewed by C Mavrot, A Kotlo; comments to author 26.07.19; revised version received 31.07.19; accepted 09.08.19; published 29.10.19.

Please cite as:
Fulton L, Kruse CS
Hospital-Based Back Surgery: Geospatial-Temporal, Explanatory, and Predictive Models
J Med Internet Res 2019;21(10):e14609
URL: http://www.jmir.org/2019/10/e14609/
doi:10.2196/14609
PMID:31663856

Viewpoint

Trust Me, I'm a Chatbot: How Artificial Intelligence in Health Care Fails the Turing Test

John Powell[1], PhD, FFPH

Nuffield Department of Primary Care Health Sciences, Medical Sciences Division, University of Oxford, Oxford, United Kingdom

Corresponding Author:
John Powell, PhD, FFPH
Nuffield Department of Primary Care Health Sciences
Medical Sciences Division
University of Oxford
Radcliffe Observatory Quarter
43 Woodstock Road
Oxford, OX2 6GG
United Kingdom
Phone: 44 1865617768 ext 617768
Fax: 44 1865289412
Email: john.powell@phc.ox.ac.uk

Abstract

Over the next decade, one issue which will dominate sociotechnical studies in health informatics is the extent to which the promise of artificial intelligence in health care will be realized, along with the social and ethical issues which accompany it. A useful thought experiment is the application of the Turing test to user-facing artificial intelligence systems in health care (such as chatbots or conversational agents). In this paper I argue that many medical decisions require value judgements and the doctor-patient relationship requires empathy and understanding to arrive at a shared decision, often handling large areas of uncertainty and balancing competing risks. Arguably, medicine requires wisdom more than intelligence, artificial or otherwise. Artificial intelligence therefore needs to supplement rather than replace medical professionals, and identifying the complementary positioning of artificial intelligence in medical consultation is a key challenge for the future. In health care, artificial intelligence needs to pass the implementation game, not the imitation game.

(J Med Internet Res 2019;21(10):e16222) doi:10.2196/16222

KEYWORDS

artificial intelligence; machine learning; medical informatics; digital health; ehealth; chatbots; conversational agents

Over the last two decades, the concerns of digital health researchers interested in the social impact of the internet have evolved as the technology has matured and new tools have emerged. From a sociotechnical perspective, there were initial preoccupations with the impact of a new, uncontrolled form of mass communication, alongside concerns with the quality of unregulated online information and threats to professions, with medical professionals in particular fearing a loss of authority [1-3]. As Web2.0 developments took hold and the public became producers as well as consumers of health information, researchers began to identify the benefits of online peer-to-peer communication and the sharing of information in virtual communities, social media, and increasingly on health ratings sites [4-7]. With the mass uptake in smartphones, the subsequent rapid developments in mobile health, and the explosion in health apps, we are now exploring the value of low-cost, patient-centered interventions delivered directly to consumers [8,9]. In addition, we are also gaining a better understanding of the limitations and key issues in their implementation, such as nonadoption and abandonment [10]. As the number one journal in this field, the Journal of Medical Internet Research continues to reflect and illuminate all these debates.

For those of us studying the social science of digital technology in health and health care, one area of research is likely to dominate the next decade: the extent to which the promise of artificial intelligence (AI) in health care will be realized, and the social and ethical issues which accompany it [11-13]. Broadly speaking, we can identify two current strands in the use of AI in health care. Firstly, there are data-facing applications which use techniques such as machine learning and artificial neural networks to derive new knowledge from large datasets, such as improving diagnostic accuracy from scans and other images [14]. Secondly, there are user-facing applications and intelligent agents which interact with people

in real-time, using inferences to provide advice or instruction based on probabilities which the tool can derive and improve over time, such as a chatbot substituting or complementing a health care consultation with a patient [15]. In this article I focus on the latter to consider the approaches of these chatbots, or "robot doctors," to medical consultation, and specifically the extent to which these technologies will ever pass the celebrated Turing test.

Alan Turing, the British mathematician and theoretical computer scientist, is widely regarded as the founding father of AI. He proposed that for a machine to be considered intelligent it should provide responses to a blinded interrogation that are indistinguishable from those given by a human comparator [16]. In other words, the interrogator should not be able to tell whether the machine or the human was responding. If we extrapolate this thought experiment to current health care, we can pose the question of whether AI-based medical consultations (conversational agents and medical chatbots) will ever be considered intelligent by Turing's standard. Of course, context is important, and if a patient is asking a simple factual question that requires a binary response, for example, then even current AI systems can mimic a human interlocutor with high accuracy. However, we know that medical consultations are complex [17], that many medical decisions require value judgements, and that the doctor-patient relationship requires empathy and understanding to arrive at a shared decision [18]. The practice of medicine is as much an art as a science, and patients may choose a path which is not necessarily the one that logic would determine. Even the pioneers of evidence-based medicine defined their normative approach as:

> *the conscientious and judicious use of current best evidence from clinical care research in the management of individual patients [19].*

Conscience and the ability to weigh competing personal values are not strengths of AI. A key skill for medical professionals is the ability to deal with uncertainty alongside considering patients' preferences. What doctors often need is wisdom rather than intelligence, and we are a long way away from a science of artificial wisdom.

It is doubtful whether AI will ever pass the Turing test for complex medical consultations, but this is to misunderstand the place of AI in future medical care. AI should complement rather than replace medical professionals. As various studies into the future of work have shown, automation in the workplace will not eliminate all human tasks [20]. Chatbot approaches have many potential benefits, including the potential to allow clinicians to have more time for delivering empathic and personalized care [15]. Perhaps, as a senior clinical informatics leader in the UK has suggested, "AI will allow doctors to be more human" [13]. However, as has been well established for many innovations in health care, especially digital ones, the key challenges for health systems seeking to harness the benefits of the technology are not just related to its effectiveness but also to the wider issues of its integration and implementation [10,12,21]. We need to understand how to integrate the tools and practices of AI within the work and culture of professionals and organizations, to investigate factors related to adoption, nonadoption, and abandonment [10,12], and investigate the work required to sustain innovation [22]. Factors which will influence the implementation of AI tools include those related to people, such as professional and public attitudes, trust, existing work practices, training needs, and the risks of deskilling and disempowerment; those related to the health system, such as leadership and management, the positioning of clinical responsibility and accountability, and the possibility of harm, alongside issues of regulation and service provision (including scalability and the possibility of providing two-tier services with or without AI); those related to the data, such as issues of data security, privacy, consent and ownership; and those related to the tool itself, such as transparency of the algorithm, issues of reliability and validity, and algorithmic bias [12,21,23]. To take an example, in an early study of an algorithm-based triage tool in primary care, we showed that physicians lacked trust in the ability of the machine to take clinical risks and worried about issues of governance and accountability, such that the sensitivity of the tool, in terms of the urgency of triage, was consistently set at a threshold which would increase urgent clinical workload rather than reduce it [24].

Identifying the complementary positioning of AI tools in health care in general, and in particular for their use in the medical consultation, is a key challenge for the future. We need to understand how to integrate the precision and power of AI tools and practices with the wisdom and empathy of the doctor-patient relationship. In health care, it is more important that artificial intelligence passes the implementation game rather than the imitation game.

Acknowledgments

JP first discussed applying the Turing test to AI in health care in 2016 and had subsequent discussions with colleagues in Oxford and elsewhere. JP is funded by the National Institute for Health Research Collaboration for Leadership in Applied Health Research and Care Oxford at Oxford Health National Health Service Foundation Trust.

Conflicts of Interest

None declared.

References

1. Hardey M. Doctor in the house: the Internet as a source of lay health knowledge and the challenge to expertise. Sociology of Health & Illness 2001 Dec 25;21(6):820-835. [doi: 10.1111/1467-9566.00185]

2. Eysenbach G, Powell J, Kuss O, Sa E. Empirical studies assessing the quality of health information for consumers on the world wide web: a systematic review. JAMA 2002 May 22;287(20):2691-2700. [doi: 10.1001/jama.287.20.2691] [Medline: 12020305]
3. Ziebland S. The importance of being expert: the quest for cancer information on the Internet. Soc Sci Med 2004 Nov;59(9):1783-1793. [doi: 10.1016/j.socscimed.2004.02.019] [Medline: 15312914]
4. Hardey M. 'E-health': the internet and the transformation of patients into consumers and producers of health knowledge. Info., Comm. & Soc 2001 Oct;4(3):388-405. [doi: 10.1080/713768551]
5. Powell J, McCarthy N, Eysenbach G. Cross-sectional survey of users of Internet depression communities. BMC Psychiatry 2003 Dec 10;3(1):1-7. [doi: 10.1186/1471-244x-3-19]
6. Eysenbach G, Powell J, Englesakis M, Rizo C, Stern A. Health related virtual communities and electronic support groups: systematic review of the effects of online peer to peer interactions. BMJ 2004 May 15;328(7449):1166 [FREE Full text] [doi: 10.1136/bmj.328.7449.1166] [Medline: 15142921]
7. van Velthoven MH, Atherton H, Powell J. A cross sectional survey of the UK public to understand use of online ratings and reviews of health services. Patient Educ Couns 2018 Sep;101(9):1690-1696 [FREE Full text] [doi: 10.1016/j.pec.2018.04.001] [Medline: 29666022]
8. Powell J, Hamborg T, Stallard N, Burls A, McSorley J, Bennett K, et al. Effectiveness of a web-based cognitive-behavioral tool to improve mental well-being in the general population: randomized controlled trial. J Med Internet Res 2012 Dec 31;15(1):e2 [FREE Full text] [doi: 10.2196/jmir.2240] [Medline: 23302475]
9. Rathbone AL, Clarry L, Prescott J. Assessing the Efficacy of Mobile Health Apps Using the Basic Principles of Cognitive Behavioral Therapy: Systematic Review. J Med Internet Res 2017 Nov 28;19(11):e399 [FREE Full text] [doi: 10.2196/jmir.8598] [Medline: 29187342]
10. Greenhalgh T, Wherton J, Papoutsi C, Lynch J, Hughes G, A'Court C, et al. Beyond Adoption: A New Framework for Theorizing and Evaluating Nonadoption, Abandonment, and Challenges to the Scale-Up, Spread, and Sustainability of Health and Care Technologies. J Med Internet Res 2017 Nov 01;19(11):e367 [FREE Full text] [doi: 10.2196/jmir.8775] [Medline: 29092808]
11. Vayena E, Blasimme A, Cohen IG. Machine learning in medicine: Addressing ethical challenges. PLoS Med 2018 Nov 6;15(11):e1002689 [FREE Full text] [doi: 10.1371/journal.pmed.1002689] [Medline: 30399149]
12. Shaw J, Rudzicz F, Jamieson T, Goldfarb A. Artificial Intelligence and the Implementation Challenge. J Med Internet Res 2019 Jul 10;21(7):e13659 [FREE Full text] [doi: 10.2196/13659] [Medline: 31293245]
13. Academy of Medical Royal Colleges. London; 2019 Jan 28. Artificial Intelligence in Healthcare URL: https://www.aomrc.org.uk/reports-guidance/artificial-intelligence-in-healthcare/ [accessed 2019-10-16]
14. Shen J, Zhang CJP, Jiang B, Chen J, Song J, Liu Z, et al. Artificial Intelligence Versus Clinicians in Disease Diagnosis: Systematic Review. JMIR Med Inform 2019 Aug 16;7(3):e10010 [FREE Full text] [doi: 10.2196/10010] [Medline: 31420959]
15. Palanica A, Flaschner P, Thommandram A, Li M, Fossat Y. Physicians' Perceptions of Chatbots in Health Care: Cross-Sectional Web-Based Survey. J Med Internet Res 2019 Apr 05;21(4):e12887 [FREE Full text] [doi: 10.2196/12887] [Medline: 30950796]
16. Turing AM. Computing Machinery and Intelligence. Mind, New Series 1950 Oct;59(236):433-460 Published by Oxford University Press on behalf of the Mind Association [FREE Full text]
17. Innes AD, Campion PD, Griffiths FE. Complex consultations and the 'edge of chaos'. Br J Gen Pract 2005 Jan;55(510):47-52 [FREE Full text] [Medline: 15667766]
18. Barry MJ, Edgman-Levitan S. Shared decision making - The pinnacle of patient-centered care. N Engl J Med 2012 Mar 01;366(9):780-781. [doi: 10.1056/NEJMp1109283] [Medline: 22375967]
19. Sackett DL, Rosenberg WMC, Gray JAM, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. BMJ 1996 Jan 13;312(7023):71-72 [FREE Full text] [doi: 10.1136/bmj.312.7023.71] [Medline: 8555924]
20. Autor DH. Why Are There Still So Many Jobs? The History and Future of Workplace Automation. Journal of Economic Perspectives 2015 Aug;29(3):3-30. [doi: 10.1257/jep.29.3.3]
21. Cresswell KM, Bates DW, Sheikh A. Ten key considerations for the successful implementation and adoption of large-scale health information technology. J Am Med Inform Assoc 2013 Jun;20(e1):e9-e13 [FREE Full text] [doi: 10.1136/amiajnl-2013-001684] [Medline: 23599226]
22. Pope C, Halford S, Turnbull J, Prichard J, Calestani M, May C. Using computer decision support systems in NHS emergency and urgent care: ethnographic study using normalisation process theory. BMC Health Serv Res 2013 Mar 23;13:111 [FREE Full text] [doi: 10.1186/1472-6963-13-111] [Medline: 23522021]
23. Greenhalgh T, Robert G, Macfarlane F, Bate P, Kyriakidou O. Diffusion of innovations in service organizations: systematic review and recommendations. Milbank Q 2004;82(4):581-629 [FREE Full text] [doi: 10.1111/j.0887-378X.2004.00325.x] [Medline: 15595944]
24. Poote AE, French DP, Dale J, Powell J. A study of automated self-assessment in a primary care student health centre setting. J Telemed Telecare 2014 Apr;20(3):123-127. [doi: 10.1177/1357633X14529246] [Medline: 24643948]

Abbreviations

AI: artificial intelligence

Edited by G Eysenbach; submitted 11.09.19; peer-reviewed by B Xie; accepted 12.10.19; published 28.10.19.

Please cite as:
Powell J
Trust Me, I'm a Chatbot: How Artificial Intelligence in Health Care Fails the Turing Test
J Med Internet Res 2019;21(10):e16222
URL: http://www.jmir.org/2019/10/e16222/
doi:10.2196/16222
PMID:31661083

Viewpoint

Role of the Internet in Solving the Last Mile Problem in Medicine

Bradford William Hesse[1], PhD

National Cancer Institute, Rockville, MD, United States

Corresponding Author:
Bradford William Hesse, PhD
National Cancer Institute
9609 Medical Center Drive
Rockville, MD, 20852
United States
Phone: 1 2402766721
Email: hesseb@mail.nih.gov

Abstract

Internet-augmented medicine has a strong role to play in ensuring that all populations benefit equally from discoveries in the medical sciences. Yet, data from the Centers for Disease Control and Prevention collected from 1999 to 2014 suggested that during the first phase of internet diffusion, progress against mortality has stalled, and in some cases, receded in rural areas that are traditionally underserved by medical and broadband resources. This problem of failing to extend the benefits of extant medical knowledge equitably to all populations regardless of geography can be framed as the "last mile problem in health care." In theory, the internet should help solve the last mile problem by making the best knowledge in the world available to everyone worldwide at a low cost and no delay. In practice, the antiquated supply chains of industrial age medicine have been slow to yield to the accelerative forces of evolving internet capacity. This failure is exacerbated by the expanding digital divide, preventing residents of isolated, geographically distant communities from taking full advantage of the digital health revolution. The result, according to the Federal Communications Commission's (FCC's) Connect2Health Task Force, is the unanticipated emergence of "double burden counties," ie, counties for which the mortality burden is high while broadband access is low. The good news is that a convergence of trends in internet-enabled health care is putting medicine within striking distance of solving the last mile problem both in the United States and globally. Specific trends to monitor over the next 25 years include (1) using community-driven approaches to bridge the digital divide, (2) addressing structural disconnects in care through P4 Medicine, (3) meeting patients at "point-of-need," (4) ensuring that no one is left behind through population management, and (5) self-correcting cybernetically through the learning health care system.

(J Med Internet Res 2019;21(10):e16385) doi:10.2196/16385

KEYWORDS

connected health; implementation science; patient engagement; community improvement; citizen science; digital divide; learning health care system

Introduction

The internet has a strong role to play in ensuring that all populations benefit equally from discoveries in the biomedical sciences. To illustrate why this is the case, consider data from the US-based Centers for Disease Control and Prevention (CDC). In 2017, the CDC revealed that despite scientific progress on multiple fronts, certain portions of the population—especially those living in isolated, nonmetropolitan areas—were falling behind across seven of the leading causes of death in the country [1]. This backward regression in mortality outcomes is especially evident in the case of cancer. From 1980 to 2014, mortality data from the cancer registries revealed that while age-adjusted mortality rates were falling in urban and suburban areas (where access to medical services and communications infrastructure was prevalent), they were atrophying or even receding in rural areas of the country. New cancer hotspots began to emerge across the country nestled within the *hollers* of Appalachia, the *bayous* of the Mississippi Delta, and the vast geographic territories covered by Native American tribal lands [2,3]. The application of extant medical knowledge was impeded by the tyranny of distance. This is not just a US phenomenon. Meta-analyses from studies conducted worldwide have revealed that in the case of cancer, people living over 50 miles away from the nearest hospital tend to present with later stages of disease, experience unaddressed complications during treatment, experience lower quality of life, and fail to comply with prescribed pharmaceutical treatments [4].

Overcoming the limitations of an inadequate, industrial age supply chain for medical knowledge can be framed as the solving the "last mile problem" in health care. The good news, according to health services researcher Don Berwick, is that the rapid diffusion of internet technologies over the past two decades is beginning to put a solution to the last mile problem within our grasp. As he indicated in his testimony to the US President's Cancer Panel, "There is now a worldwide collection of efforts, which is showing how much we can leverage knowledge through (health information technology) so that literally the best knowledge in the world can reach everyone in the world, at low cast, and at no delay." [5]. This is the promise that those of us working at the intersection of medicine and the internet can help fulfil with vision and collaboration. The objective is worth our collective efforts. With respect to the evocative case of cancer, the American Cancer Society estimated that solving the access issue in oncology would improve mortality rates by approximately 22% per year [6].

From Telemedicine to Connected Health: The Last 25 Years

In its earliest days, the idea of reaching remote populations through electronic means—a concept embodied in the practice of telemedicine—seemed to be a hopeful, but often an impracticable, solution for solving the last mile problem. Video-conferencing capabilities were expensive and technologically challenging, electronic health record (EHR) systems were rare, and patients' abilities to reach their clinical teams through electronic communication were practically nonexistent. The world's medical knowledge was still locked up in the stacks of academic medical libraries, which were often nonexistent in rural settings and inaccessible to practitioners in low-resourced countries. Patients were generally precluded from accessing medical knowledge directly, because they lacked access to professional distribution channels and the material was written in a highly specialized medical language that was inaccessible to anyone without a medical education. These industrial-age dissemination channels forced reliance on highly trained clinical personnel as mediators of medical knowledge and awarded a premium to the elite medical schools as purveyors of evidence.

In the early 1990s, the US-based National Science Foundation (NSF) invested in a strategy that would overcome the limitations of geography in science by connecting the world's knowledge resources through a hyperlinked lattice of documents, data, remote devices, and personal communications. Under the leadership of Donald Lindberg, the National Library of Medicine joined the NSF in its vision for accessible online knowledge by digitizing its holdings and making them available in its online bibliographic resource, MEDLINE. The objective of MEDLINE was to help remote practitioners overcome the limitations of underresourced libraries and to help all communities benefit from an up-to-date, comprehensive snapshot of the extant medical knowledge base. Soon afterward, legislative and regulatory bodies made a set of decisions that would open the internet to the general public. Public-facing search engines directed anyone with a computer and a modem to the same bibliographic databases being used by their doctors. Biomedical visionaries predicted that electronic health would soon take its place along with electronic commerce as the new distribution channel for medical knowledge in health generally [7-9] and cancer specifically [10,11].

As the internet matured, it entered the mainstream of commercial, civic, and social life with substantial repercussions throughout. Initial forays into electronic commerce stumbled, bringing about the dot.com implosion at the beginning of the millennium, but as companies returned to the first principles, the practice flourished. In medicine, demand seemed to precede supply as patients flocked to the Web first, even before visiting their doctors, in search of reliable guidance on how to care for themselves or their families [12]. Patients' online reconnaissance was not always appreciated by the medical establishment, which had been resisting the digitization of its own internal processes. In the absence of receptivity from the medical establishment, patients found each other online, exchanging insights through bulletin boards, chat rooms, and eventually social media. "Health 2.0" took on the form of a grassroots effort to encourage greater acceptance of engagement and innovation by patients across the supply chain [13]. In the United States, it would (literally) take an act of Congress to create the incentives for hospital systems and individual physicians to move from paper-based records to EHRs [14]. Even then, the early functionality of these EHRs would be oriented primarily toward billing and coding purposes. They were not designed with the appropriate human factors expertise to ease pressures on workflow [15] or to empower patients [16].

Now, 10 years after passage of the Health Information Technology for Economic and Clinical Health Act [17] in the United States, the conditions appear to be in place for substantial innovation to improve the medical knowledge supply chain. Adoption of EHRs in medicine has reached an all-time high, with 96% of nonfederal acute care hospitals and 86% of office-based physicians attesting to the meaningful use of health information technology [18]. Usability and safety issues are no longer swept under the table but have taken on a more centralized role in contemporary legislation [19]. Access to smartphone technology skyrocketed after the introduction of the iPhone in 2007, giving patients always-on, always-present access to the functionality of internet-based channels [20]. The wearable device and medical sensor markets have also been growing, giving medical entrepreneurs an opportunity to extend care more seamlessly into the home. The recent entry of a major device and software manufacturer into the personal health record market prompted the US General Accounting Office to declare that the market around consumer engagement in medicine may have finally reached a tipping point.

Solving the Last Mile Problem: The Next 25 Years

Despite a tortuous path in bringing medicine into the dawn of the digital age, the conditions are now in place to make exponential progress in solving the last mile problem in health care. The following are some of the areas worthy of emphasis and continued monitoring over the next 25 years:

Using Community-Driven Approaches to Bridge the Digital Divide

In the United States, the FCC's Connect2Health Task Force has been monitoring the extent to which broadband capacity is available throughout the country to support the digitization of health care. Paradoxically, many of the same rural counties experiencing a decline in progress against the biggest threats to mortality according to the CDC are also falling outside the reach of reliable broadband coverage. These counties are experiencing a dual burden: The mortality burden from chronic disease is higher in these counties than in their metropolitan counterparts, while the infrastructure to reach these patients through community hospitals or, now telemedicine, is dwindling. In 2017, the FCC and the National Cancer Institute initiated a broadband pilot program in rural Appalachia called "Linking and Amplifying User-Centered Networks through Connected Health," or LAUNCH. The program is using expertise from the Design Lab at the University of California San Diego; expertise for value-based partnerships from Amgen; oncology expertise from the Markey Cancer Center in Lexington, Kentucky; and channeled expertise from industry leaders to create a platform for community-driven development on top of an extended platform for connected services through broadband [21].

Addressing Structural Disconnects in Care Through P4 Medicine

Twentieth century medicine bore the hallmarks of the industrial age, with one-size-fits-all solutions dominating the marketplace as blockbuster products and fee-for-service treatment centers offering the promise of indemnified cures, or repairs, after symptoms became too bothersome to ignore. The trouble was that these solutions, which were reactionary in nature and delivered too late in the disease process to prevent irreparable damage, were insufficient to cope with the deluge of noncommunicable diseases projected to drain the coffers of social support systems internationally. Leaders in medicine have called for a new approach, one enabled by the 21st century internet technologies. The new approach must be predictive, using cutting edge analytics to forecast risk; preemptive, utilizing prevention and early detection formulae to act upon those risk profiles before damage occurs; personalized, tailoring treatments to patients' risk and preference profiles to create solutions that are both efficacious and value congruent; and participative, embracing the collaborative capacity of internet platforms to support patient engagement, community improvement, and citizen science [22]. To enable this new brand of medicine (referred to by some as "P4 Medicine"), health system engineers must focus on eliminating the disconnects in care that have otherwise derailed efficacious, empowered action across care teams inclusive of patients and their caregivers [23]. Just as FedEx and Amazon dominated the marketplace by ensuring custodial responsibility across the consumer fulfillment supply chain, new leaders in biomedicine will be those who eliminate the structural disconnects in care to ensure custodial support for covered lives across the interconnected supply chains delivering preemptive and participative care.

Meeting Patients at the Point of Need

Just as internet-supported medicine can be mobilized to solve the last mile problem in terms of geography, data suggest that it can also be used to solve access issues as imposed by temporal constraints. Twentieth century medicine was dominated by the office visit or hospital stay, the clinical equivalent of bricks-and-mortar service delivery. The problem is that most patients' health decisions occur outside of the clinical context [24]. NCI-funded clinical trials have already demonstrated just how effective the use of asynchronously collected patient-reported outcomes can be in helping cancer patients stave off the unanticipated side effects of treatment while reducing costs from preventable hospitalizations or controllable symptomologies. New efforts are underway within the biomedical technology sector to create the next generation of biologic sensors that can be used to place an electronic safety net around patients when they are away from the clinic. Usability engineers are designing more patient-centric ways for patients to ask questions and manage their care asynchronously while on travel or at home. Advanced medical sensors are being developed using nanotechnology (eg, microneedle sensors) to reduce the obtrusiveness of physiologic monitoring, while connected smart devices in the home are being engineered to detect deviations in air quality, falls, or complications from treatment passively and unobtrusively [25,26]. Tracking and understanding how these internet-connected devices can be used to support better patient outcomes through ongoing support outside of the home, while protecting privacy and safety, was a priority embedded within the 21st Century Cures Act passed in 2016 [27]. Further integrating the data from these devices into a supportive platform for personal care management and remote clinical monitoring without overwhelming the health care system will be the human factors challenge for the next 25 years.

Ensuring That Nobody is Left Behind Through Population Management

When meaningful use requirements were first articulated for the incentives intended to spur adoption of Health Information Technology in the United States in 2009, one essential policy lever was included that pointed to a dramatic restructuring of the way medical care is delivered. The lever was population health management. Its purpose was to provide health insurers and health care providers the ability to go beyond individual patient management to a level of proactively managing the health and welfare of all patients within a specific practice, or to ensure that all patients are served equitably within the population of a health care plans' members. Dr Nirav Shah, who was the Vice President for Kaiser Permanente in Southern California, delivered a poignant example of successful population health management in testimony to the President's Cancer Panel in the spring of 2015. In his example, Kaiser used the tracking capacity of its mature EHR system to monitor patients' recommended eligibility for age-/risk-based cancer screening. Office staff proactively followed prompts from the system to ensure that everyone on the list had been contacted with a recommendation for the screening. Results showed a 6-fold increase in mammogram testing, a 6-fold increase in cervical cancer testing, and a 10-fold increase in colorectal

cancer testing. External reviews revealed an absence of disparate outcomes for the screening exercise; the approach benefitted all population with the membership equitably [5].

Self-Correcting Cybernetically Through the Learning Health Care System

The other quantum leap forward enabled by connective technology is the premise that data volunteered by patients and harvested from administrative systems can be used to improve the performance of health care. Many of the largest consumer-facing electronics systems routinely request consent to gather data on reported system errors to improve the fidelity of their products in the field. In the era of internet-connected things, usage data from passive sensors can be used to adjust load levels at the community level, while data from personal voice assistants can be fed into machine-learning algorithms to improve sensitivity and performance of automated services. PatientsLikeMe.com, billed as the world's largest personalized health network, helped spark a revolution in biomedicine by bringing citizen science to bear on the development and postmarket monitoring of life-saving therapeutics. The National Institutes of Health plunged into this revolution as it launched the All of Us initiative, a program designed to bring volunteered data from a million-patient cohort directly into the discovery engines of biomedicine. Taken together, these emerging capacities provide an early vision for how patients, providers, and scientists can participate together to realize what the National Academy of Medicine has referred to as a true Learning Healthcare System [28-30].

Conclusion and Caveat

Data from the CDC offer a poignant reminder to the limits of industrial age medicine. Traditional supply chains are slow, expensive, and restrictive in their abilities to translate the benefits of hard-won medical knowledge equitably to all patients, regardless of where they live or what time constraints govern their days. As a result, broad swaths of the world's population are being left behind, receiving neither the benefits of evidence-based knowledge nor the opportunity to participate equitably in the discovery of tomorrow's cures. This is the last mile problem in medicine. A new paradigm, enabled by internet technology, brings hope that medicine can work collectively to solve the last mile problem over the next 25 years.

This hope, however, comes with a caveat. The same conditions that gave patients direct access to the scientific literature otherwise sequestered in the world's most prestigious libraries, have also given rise to a social milieu in which medical misinformation can spread as quickly as medical fact [31]. Similarly, the same technology that allows for precision cataloging of a patient's personal genome to be considered in tandem with biologic data from implanted sensors or contextual data from wearable device is rapidly creating an explosion of data, which, if left untethered, may prove to be paralyzing to decision makers [32]. Hard work will be needed to track the unanticipated consequences of exponential disruption in the medical space and then use the best principles of human-centered design to address them directly for the benefit of patient outcomes and public health. I am heartened to know that at that point, we will be able to read about the fruits of this labor in the publications of the Journal of Medical Internet Research.

Conflicts of Interest

None declared.

References

1. Moy E, Garcia MC, Bastian B, Rossen LM, Ingram DD, Faul M, et al. Leading Causes of Death in Nonmetropolitan and Metropolitan Areas - United States, 1999-2014. MMWR Surveill Summ 2017 Jan 13;66(1):1-8 [FREE Full text] [doi: 10.15585/mmwr.ss6601a1] [Medline: 28081058]
2. Mokdad AH, Dwyer-Lindgren L, Fitzmaurice C, Stubbs RW, Bertozzi-Villa A, Morozoff C, et al. Trends and Patterns of Disparities in Cancer Mortality Among US Counties, 1980-2014. JAMA 2017 Jan 24;317(4):388-406 [FREE Full text] [doi: 10.1001/jama.2016.20324] [Medline: 28118455]
3. Wheeler SB, Basch E. Translating Cancer Surveillance Data Into Effective Public Health Interventions. JAMA 2017 Jan 24;317(4):365-367. [doi: 10.1001/jama.2016.20326] [Medline: 28118434]
4. Ambroggi M, Biasini C, Del Giovane C, Fornari F, Cavanna L. Distance as a Barrier to Cancer Diagnosis and Treatment: Review of the Literature. Oncologist 2015 Dec;20(12):1378-1385 [FREE Full text] [doi: 10.1634/theoncologist.2015-0110] [Medline: 26512045]
5. Improving Cancer-Related Outcomes with Connected Health. Washington, DC: President's Cancer Panel; 2016 Oct. URL: https://prescancerpanel.cancer.gov/report/connectedhealth/pdf/PresCancerPanel_ConnHealth_Nov2016.pdf
6. Siegel RL, Jemal A, Wender RC, Gansler T, Ma J, Brawley OW. An assessment of progress in cancer control. CA Cancer J Clin 2018 Sep;68(5):329-339 [FREE Full text] [doi: 10.3322/caac.21460] [Medline: 30191964]
7. Eysenbach G. What is e-health? J Med Internet Res 2001 Jun 18;3(2):E20 [FREE Full text] [doi: 10.2196/jmir.3.2.e20] [Medline: 11720962]
8. Atienza AA, Hesse BW, Baker TB, Abrams DB, Rimer BK, Croyle RT, et al. Critical issues in eHealth research. Am J Prev Med 2007 May;32(5 Suppl):S71-S74 [FREE Full text] [doi: 10.1016/j.amepre.2007.02.013] [Medline: 17466821]

9. Gibbons MC. eHealth solutions for healthcare disparities. New York, London: Springer; 2007.
10. Eysenbach G. The impact of the Internet on cancer outcomes. CA Cancer J Clin 2003;53(6):356-371 [FREE Full text] [doi: 10.3322/canjclin.53.6.356] [Medline: 15224975]
11. Hesse BW. Harnessing the power of an intelligent health environment in cancer control. Stud Health Technol Inform 2005;118:159-176. [Medline: 16301777]
12. Mohr M, Henche HR. The morbidity associated with lost or irretrievable resected meniscal fragments. Arthroscopy: The Journal of Arthroscopic & Related Surgery 1992 Mar;8(1):84-88. [doi: 10.1016/0749-8063(92)90140-7] [Medline: 1550655]
13. O'Reilly T. O'Reilly. 2005 Sep 30. What Is Web 2.0? Design Patterns and Business Models for the Next Generation of Software URL: http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html
14. Blumenthal D. Launching HITECH. N Engl J Med 2010 Feb 04;362(5):382-385. [doi: 10.1056/NEJMp0912825] [Medline: 20042745]
15. Tevaarwerk AJ, Klemp JR, van Londen GJ, Hesse BW, Sesto ME. Moving beyond static survivorship care plans: A systems engineering approach to population health management for cancer survivors. Cancer 2018 Nov 15;124(22):4292-4300. [doi: 10.1002/cncr.31546] [Medline: 30277575]
16. Hesse BW, Shneiderman B. eHealth research from the user's perspective. Am J Prev Med 2007 May;32(5 Suppl):S97-103 [FREE Full text] [doi: 10.1016/j.amepre.2007.01.019] [Medline: 17466825]
17. 111th Congress of U.S. HITECH Act. Representatives USSaHo, ed. 42 USC 139w-(2). USA: US House of Representatives; 2009. URL: https://uscode.house.gov/statutes/pl/111/232.pdf [accessed 2019-10-16]
18. The Office of the National Coordinator for Health Information Technology. 2017. Health IT Dashboard: Quick Stats URL: https://dashboard.healthit.gov/quickstats/quickstats.php [accessed 2019-10-16]
19. Ratwani RM, Hodgkins M, Bates DW. Improving Electronic Health Record Usability and Safety Requires Transparency. JAMA 2018 Dec 25;320(24):2533-2534. [doi: 10.1001/jama.2018.14079] [Medline: 30489619]
20. Serrano KJ, Thai CL, Greenberg AJ, Blake KD, Moser RP, Hesse BW. Progress on Broadband Access to the Internet and Use of Mobile Devices in the United States. Public Health Rep 2017;132(1):27-31 [FREE Full text] [doi: 10.1177/0033354916679365] [Medline: 28005473]
21. Federal Communications Commission. FCC-NCI Broadband Cancer Collaboration URL: https://www.fcc.gov/health/cancer [accessed 2019-10-16]
22. Shaikh AR, Butte AJ, Schully SD, Dalton WS, Khoury MJ, Hesse BW. Collaborative biomedicine in the age of big data: the case of cancer. J Med Internet Res 2014 Apr 07;16(4):e101 [FREE Full text] [doi: 10.2196/jmir.2496] [Medline: 24711045]
23. Hesse B, Ahern BW, Beckjord D, Ellen B. Oncology Informatics: Using Health Information Technology to Improve Processes and Outcomes in Cancer. Boston, MA: Academic Press; 1969.
24. Dentzer S. Health Care Without Walls: A Roadmap for Reinventing U.S. Health Care. Washington, DC: Network for Excellence in Health Innovation; 2018.
25. Kvedar JC, Colman C, Cella G. The Internet of Healthy Things. Boston, MA: The Internet of Healthy Things; 2015.
26. Kvedar JC, Colman C, Cella G. The New Mobile Age: How Technology Will Extend The Healthspan And Optimize The Lifespan. Boston, MA: The Internet Of Healthy Things; 2017.
27. Congress.gov. Washington, DC: US Congress; 2016. One Hundred Fourteenth Congress of the United States of America URL: https://www.congress.gov/114/bills/hr34/BILLS-114hr34enr.pdf [accessed 2019-10-16]
28. Institute of Medicine (US) and National Academy of Engineering (US) Roundtable on Value & Science-Driven Health Care. Engineering a Learning Healthcare System: A Look at the Future: Workshop Summary. Washington, DC: National Academies Press (US); 2011. URL: https://www.ncbi.nlm.nih.gov/books/NBK50934/
29. Schilsky RL, Michels DL, Kearbey AH, Yu PP, Hudis CA. Building a rapid learning health care system for oncology: the regulatory framework of CancerLinQ. J Clin Oncol 2014 Aug 01;32(22):2373-2379. [doi: 10.1200/JCO.2014.56.2124] [Medline: 24912897]
30. Institute of Medicine, Committee on the Learning Health Care System in America, Stuckhardt L, Saunders R, Smith M. Best Care At Lower Cost: The Path To Continuously Learning Health Care In America. Washington, DC: National Academies Press; 2019.
31. Chou WS, Oh A, Klein WMP. Addressing Health-Related Misinformation on Social Media. JAMA 2018 Dec 18;320(23):2417-2418. [doi: 10.1001/jama.2018.16865] [Medline: 30428002]
32. Wachter R. The Digital Doctor: Hope, Hype, and Harm at the Dawn of Medicine's Computer Age. New York, NY: Mcgraw-Hill Education; 2019.

Abbreviations

CDC: Centers for Disease Control and Prevention
EHR: electronic health records
FCC: Federal Communications Commission’s
NSF: National Science Foundation

Edited by G Eysenbach; submitted 24.09.19; peer-reviewed by S Chen; accepted 12.10.19; published 28.10.19.

Please cite as:
Hesse BW
Role of the Internet in Solving the Last Mile Problem in Medicine
J Med Internet Res 2019;21(10):e16385
URL: http://www.jmir.org/2019/10/e16385/
doi:10.2196/16385
PMID:31661078

Viewpoint

The Importance of Health Information on the Internet: How It Saved My Life and How it Can Save Yours

Andre Kushniruk[1], MSc, PhD

School of Health Information Science, University of Victoria, Victoria, BC, Canada

Corresponding Author:
Andre Kushniruk, MSc, PhD
School of Health Information Science
University of Victoria
Human and Social Development Building room A202
3800 Finnerty Road (Ring Road)
Victoria, BC, V8P5C2
Canada
Phone: 1 250 472 5132
Fax: 1 250 472 4751
Email: andrek@uvic.ca

Abstract

The internet holds the potential promise of improved patient outcomes, especially when one is faced with a critical or life-threatening disease or condition. Appropriate and timely access to health information can support informed negotiation of optimal treatments, optimal management, and expedited recovery, and to an improved outcome for a patient. However, there are many human and technical barriers that may prevent the application of the best possible information for both patient and provider alike, making the patient journey complex and potentially dangerous. In this viewpoint paper, the author (who is also a JMIR editor) reflects on a personal patient journey, where use of the internet facilitated a means of reaching a good patient outcome in the face of a variety of informational and organizational limitations and gaps. This journey illustrates the importance of human-related factors affecting access to health information. The application of a range of internet information resources at critical points can result in a positive patient outcome, as this case illustrates. This paper reflects on how the experience highlights several information needs and concerns. It also highlights the need for improved access to appropriate health information along the patient journey that can support patient and provider joint decision-making. This access to information can make the difference between positive clinical outcomes and death, illustrating how health information on the internet can be both critical and life saving.

(J Med Internet Res 2019;21(10):e16690) doi:10.2196/16690

KEYWORDS

patient journey; human factors; consumer health informatics; eHealth; digital health; participatory medicine; shared decision-making; cancer information; tongue cancer

Introduction

The Journal of Medical Internet Research (JMIR) is now celebrating its 20th anniversary and has become a leading journal in the area of health informatics and digital health, particularly in the area of using the internet to provide health information to health care providers, patients, and average citizens alike. The importance of the focus of the journal (and its sister JMIR journals) cannot be overstated, which I especially recognize due to a recent health experience that I would like to relate in this viewpoint paper. I believe it has deep meaning for me, but I also believe that it has implications for the health care system and patient interactions with it. My case will hopefully highlight some of the issues related to availability of information for patients, access to the best possible information, and how that information needs to be used and acted on to lead to the best possible patient health outcomes. In this story, the central theme is the need for access to relevant health information on the internet and translating that information into the best possible care.

My Patient Journey

I am the editor-in-chief of JMIR Human Factors, a health informatics specialist (working both in academia and in industry in Canada and worldwide) and have been working in the field for over 25 years, but had never been on the other side of the health care system as a patient. This all changed dramatically

for me in the fall of 2017 when I was diagnosed with advanced stage tongue cancer. The cancer had been missed in the preceding summer by two dentists (my condition was attributed to other, more minor problems, as it was judged unlikely to be cancer because I had none of the risk factors). By the time I reached my general practitioner, I had realized there was something more serious behind my symptoms and I entered the waiting phase (approximately a month at that point) to see an ear, nose, and throat (ENT) specialist on what was supposed to be an urgent visit. While waiting for this appointment I began using the internet to explore the possible causes of my problem, as my symptoms worsened rapidly. After waiting several weeks to see the specialist in my local area, the problem had become exacerbated. At that time, I was told I had a tumor and would require a biopsy and imaging results before I could proceed to treatment (which took several additional weeks in total to get). This was shocking news for me, but it was nowhere near as shocking as when a few weeks later the same ENT surgeon told me that the tumor was too large to operate on. As a result, I would not be receiving life-saving surgery and as a consequence I was now considered to be palliative, with a very slim statistical chance of surviving past two years.

Fortunately, my wife (an accomplished health informatician and also editor-in-chief of JMIR Nursing) decided to apply her knowledge and skill in searching the internet to examine the assessment of the specialist. Using her skills in searching PubMed, Google, YouTube, and patient blogs on the internet, she was able to determine that the situation was not completely irreversible; in other localities life-saving surgery was possible and had been conducted for patient cases very similar to mine (and reported on in the literature of peer reviewed articles). Indeed, a search of the American and Canadian cancer society websites also indicated that there was hope and that my survival chances should have been at least double what I was initially told. Through careful search of the leading articles on PubMed, we were able to locate a hospital where I would have the best chances of being cured. Coincidentally, the hospital was in New York, was a major international center, and turned out to be the part of the same organization where I had worked as an adjunct professor as far back as 2001. At that point I contacted my colleagues in New York, and within three days I was able to get an appointment with one of the top surgeons in the world and I was able to have life-saving surgery only a week later. This surgery consisted of a glossectomy and neck dissection, followed by a tongue reconstruction using leg muscle. From my first visit we discussed treatment options and I felt very comfortable, especially as all that I was told by the specialist in New York was consistent with the latest evidence available over the internet and on Medline (his group had actually generated much of the seminal research publications in the area). I was able to receive the completely successful surgery, but I did require months of recovery, follow-up radiation and chemotherapy, which I received once back in Canada.

When I returned to Canada, the focus of my attention shifted from a curative focus to how to recover from extensive surgery, chemotherapy, and radiation to regain my ability to eat and speak. Here my wife and I accessed further articles from PubMed, explored blogs and postings on YouTube and Pinterest, and posts by patients who had undergone similar treatment and therapy. This hugely helped in not only giving me realistic expectations regarding recovery (which I could not fully get from health professionals alone) but also in sharing strategies and tips for the road to full recovery. Once again, the internet became a big part of my patient journey.

I am now two years cancer free and resumed all the activities I undertook before becoming ill. Much of what happened, and how I was able to turn around what appeared to be a bleak situation, I can attribute to my friends and colleagues who helped me get the relevant information from the internet that helped with decision making about my health and well-being.

Implications

Over the past two years I have reflected on what I have learned from this experience (as this is something I have been trained to do). My experience pointed to a variety of information gaps, issues from my human perspective, and the need for accessible, credible, and evidence-based information. Putting a human factors lens to my story, I have tried to extrapolate several lessons from my patient journey and from that of other patients I encountered during this period:

- A need for accessible information for patients on the internet (accessing this information was life saving for me) [1].
- A need for knowledge about how to access the right information. This is the advantage I had of being a specialist, but it should be made more accessible through improved user interfaces for those who do not have a background in health care [2].
- A need for access to the best possible care and treatment plans, and the ability to identify the most appropriate and best physicians available. Here the internet was what led me to locate a surgeon capable of turning my situation around [3].
- Improved electronic health (eHealth) literacy, to help in integrating technological skills with patient reasoning about critical health conditions [4,5].
- A need for credible, up to date, and substantiated evidence-based information from anywhere in the world [6].
- A need for new systems and technologies to speed up wait time and diagnosis, and to obtain second opinions (eg, easily accessible virtual second opinion systems) [7].
- A need for patients to be more informed about choices and statistics, including the meaning of survival curves in relation to different treatment options [2].
- A need for patients to be able to critique different treatment options and be provided independent advocacy and support in doing so [2].
- A need for patient education about how to select reliable and reputable information sources, requiring that information from YouTube and other such sources be curated or vetted to be up to date and useful for patient decision making [8].
- A need for integration of information and expertise, whether physical or technological (eg, a virtual tumor discussion board would have been helpful in my case) [9].

- A need for information about, and access to, life-saving treatment methods that may not be available in a patient's local area [7].
- A need for patients to continue to provide support and advice to other patients over the internet using social media and virtual communities [3].

In conclusion, I feel my story and patient journey highlight the need for usable and useful interfaces that allow the public to access the right health information, resources, specialists, and treatment in a timely manner. Reflecting on my experience, the role of eHealth and consumer informatics for ensuring evidence-based patient choice cannot be overstated. Indeed, my patient journey highlights the need for evidence-based patient choice, which has the requirements that:

> *objective, unbiased information must be made available to the patient, and the patient must have the power and opportunity to choose [10].*

However, the information used to support shared decision-making between the patient and provider must be of high quality. Only with timely access to the most appropriate information, along with the opportunity to select an appropriate decision path based on that information, can patients become truly empowered [11]. In line with the effort to improve the quality, access, and use of health information, I am proud and honored to be involved as an editor with JMIR, the leading journal for promoting the dissemination of scientific and evidence-based studies, methods, and approaches to improving health for patients like myself. And yes, health information on the internet can literally save your life, a message that has really resonated with me over the past two years.

Conflicts of Interest

None declared.

References

1. Eysenbach G, Köhler C. How do consumers search for and appraise health information on the world wide web? Qualitative study using focus groups, usability tests, and in-depth interviews. BMJ 2002 Mar 09;324(7337):573-577 [FREE Full text] [doi: 10.1136/bmj.324.7337.573] [Medline: 11884321]
2. Wills CE, Holmes-Rovner M. Patient comprehension of information for shared treatment decision making: state of the art and future directions. Patient Education and Counseling 2003 Jul;50(3):285-290. [doi: 10.1016/s0738-3991(03)00051-x]
3. Eysenbach G. The impact of the Internet on cancer outcomes. CA Cancer J Clin 2003 Nov 01;53(6):356-371 [FREE Full text] [doi: 10.3322/canjclin.53.6.356] [Medline: 15224975]
4. Kayser L, Kushniruk A, Osborne RH, Norgaard O, Turner P. Enhancing the Effectiveness of Consumer-Focused Health Information Technology Systems Through eHealth Literacy: A Framework for Understanding Users' Needs. JMIR Hum Factors 2015 May 20;2(1):e9 [FREE Full text] [doi: 10.2196/humanfactors.3696] [Medline: 27025228]
5. Monkman H, Kushniruk A. eHealth literacy issues, constructs, models, and methods for health information technology design and evaluation. Knowledge Management & E-Learning: An International Journal 2015;7(4):541-549. [doi: 10.34105/j.kmel.2015.07.036]
6. Kobayashi R, Ishizaki M. Examining the Interaction Between Medical Information Seeking Online and Understanding: Exploratory Study. JMIR Cancer 2019 Sep 24;5(2):e13240 [FREE Full text] [doi: 10.2196/13240] [Medline: 31586366]
7. Santiago TC, Jenkins JJ, Pedrosa F, Billups C, Quintana Y, Ribeiro RC, et al. Improving the histopathologic diagnosis of pediatric malignancies in a low-resource setting by combining focused training and telepathology strategies. Pediatr Blood Cancer 2012 Aug 07;59(2):221-225. [doi: 10.1002/pbc.24071] [Medline: 22315236]
8. Tan M, Kok K, Ganesh V, Thomas S. Patient information on breast reconstruction in the era of the world wide web. A snapshot analysis of information available on youtube.com. Breast 2014 Feb;23(1):33-37 [FREE Full text] [doi: 10.1016/j.breast.2013.10.003] [Medline: 24215982]
9. Marshall CL, Petersen NJ, Naik AD, Vander Velde N, Artinyan A, Albo D, et al. Implementation of a regional virtual tumor board: a prospective study evaluating feasibility and provider acceptance. Telemed J E Health 2014 Aug;20(8):705-711 [FREE Full text] [doi: 10.1089/tmj.2013.0320] [Medline: 24845366]
10. Eysenbach G, Diepgen TL. The role of e-health and consumer health informatics for evidence-based patient choice in the 21st century. Clinics in Dermatology 2001 Jan;19(1):11-17. [doi: 10.1016/s0738-081x(00)00202-9]
11. Coulter A, Entwistle V, Gilbert D. Sharing decisions with patients: is the information good enough? BMJ 1999 Jan 30;318(7179):318-322 [FREE Full text] [doi: 10.1136/bmj.318.7179.318] [Medline: 9924064]

Abbreviations

eHealth: electronic health
ENT: ears, nose, and throat
JMIR: Journal of Medical Internet Research

Edited by G Eysenbach; submitted 14.10.19; peer-reviewed by D Protti; accepted 18.10.19; published 27.10.19.

Please cite as:
Kushniruk A
The Importance of Health Information on the Internet: How It Saved My Life and How it Can Save Yours
J Med Internet Res 2019;21(10):e16690
URL: http://www.jmir.org/2019/10/e16690/
doi:10.2196/16690
PMID:31661084

White Paper

An Integrated Brain-Machine Interface Platform With Thousands of Channels

Elon Musk[1]; Neuralink[1]

Neuralink, San Francisco, CA, United States

Corresponding Author:
Elon Musk
Neuralink
3180 28th Street
San Francisco, CA, 94110
United States
Phone: 1 800 123 4567

Related Articles:

Comment in: http://jmir.org/2019/10/e16321/

Comment in: http://jmir.org/2019/10/e16339/

Comment in: http://jmir.org/2019/10/e16356/

Comment in: http://jmir.org/2019/11/e16344/

Abstract

Brain-machine interfaces hold promise for the restoration of sensory and motor function and the treatment of neurological disorders, but clinical brain-machine interfaces have not yet been widely adopted, in part, because modest channel counts have limited their potential. In this white paper, we describe Neuralink's first steps toward a scalable high-bandwidth brain-machine interface system. We have built arrays of small and flexible electrode "threads," with as many as 3072 electrodes per array distributed across 96 threads. We have also built a neurosurgical robot capable of inserting six threads (192 electrodes) per minute. Each thread can be individually inserted into the brain with micron precision for avoidance of surface vasculature and targeting specific brain regions. The electrode array is packaged into a small implantable device that contains custom chips for low-power on-board amplification and digitization: The package for 3072 channels occupies less than 23×18.5×2 mm^3. A single USB-C cable provides full-bandwidth data streaming from the device, recording from all channels simultaneously. This system has achieved a spiking yield of up to 70% in chronically implanted electrodes. Neuralink's approach to brain-machine interface has unprecedented packaging density and scalability in a clinically relevant package.

(J Med Internet Res 2019;21(10):e16194) doi:10.2196/16194

KEYWORDS

brain-machine interface; sensory function; motor function; neurology

Introduction

Brain-machine interfaces have the potential to help people with a wide range of clinical disorders. For example, researchers have demonstrated human neuroprosthetic control of computer cursors [1-3], robotic limbs [4,5], and speech synthesizers [6] by using no more than 256 electrodes. Although these successes suggest that high-fidelity information transfer between brains and machines is possible, development of brain-machine interface has been critically limited by the inability to record from large numbers of neurons. Noninvasive approaches can record the average of millions of neurons through the skull, but this signal is distorted and nonspecific [7,8]. Invasive electrodes placed on the surface of the cortex can record useful signals, but they are limited in that they average the activity of thousands of neurons and cannot record signals deep in the brain [9]. Most brain-machine interfaces have used invasive techniques, because the most precise readout of neural representations requires

recording single action potentials from neurons in distributed, functionally linked ensembles [10].

Microelectrodes are the gold-standard technology for recording action potentials, but there is no clinically translatable microelectrode technology for large-scale recordings [11]. This would require a system with material properties that provide high biocompatibility, safety, and longevity. Moreover, this device would also need a practical surgical approach and high-density, low-power electronics to ultimately facilitate fully implanted wireless operation.

Most devices for long-term neural recording are arrays of electrodes made from rigid metals or semiconductors [12-18]. Although rigid metal arrays facilitate penetrating the brain, the size, Young modulus, and bending stiffness mismatches between stiff probes and brain tissue can drive immune responses that limit the function and longevity of these devices [19,11]. Furthermore, the fixed geometry of these arrays constrains the populations of neurons that can be accessed, especially due to the presence of vasculature.

An alternative approach is to use thin, flexible multielectrode polymer probes [20,21]. The smaller size and increased flexibility of these probes should offer greater biocompatibility. However, a drawback of this approach is that thin polymer probes are not stiff enough to directly insert into the brain; their insertion must be facilitated by stiffeners [22,21], injection [23,24], or other approaches [25], all of which are quite slow [26,27]. To satisfy the functional requirements for a high-bandwidth brain-machine interface, while taking advantage of the properties of thin-film devices, we developed a robotic approach, where large numbers of fine and flexible polymer probes are efficiently and independently inserted across multiple brain regions [28].

Here, we report Neuralink's progress toward a flexible, scalable brain-machine interface that increases channel count by an order of magnitude over prior work. Our system has three main components: ultra-fine polymer probes, a neurosurgical robot, and custom high-density electronics (all of which are described below). We demonstrate the rapid implantation of 96 polymer threads, each thread with 32 electrodes, yielding a total of 3072 electrodes.

We developed miniaturized custom electronics that allow us to stream full broadband electrophysiology data simultaneously from all these electrodes (described below). We packaged this system for long-term implantation and developed custom online spike-detection software that can detect action potentials with low latency. Together, this system serves as a state-of-the-art research platform and the first prototype toward a fully implantable human brain-machine interface.

Threads

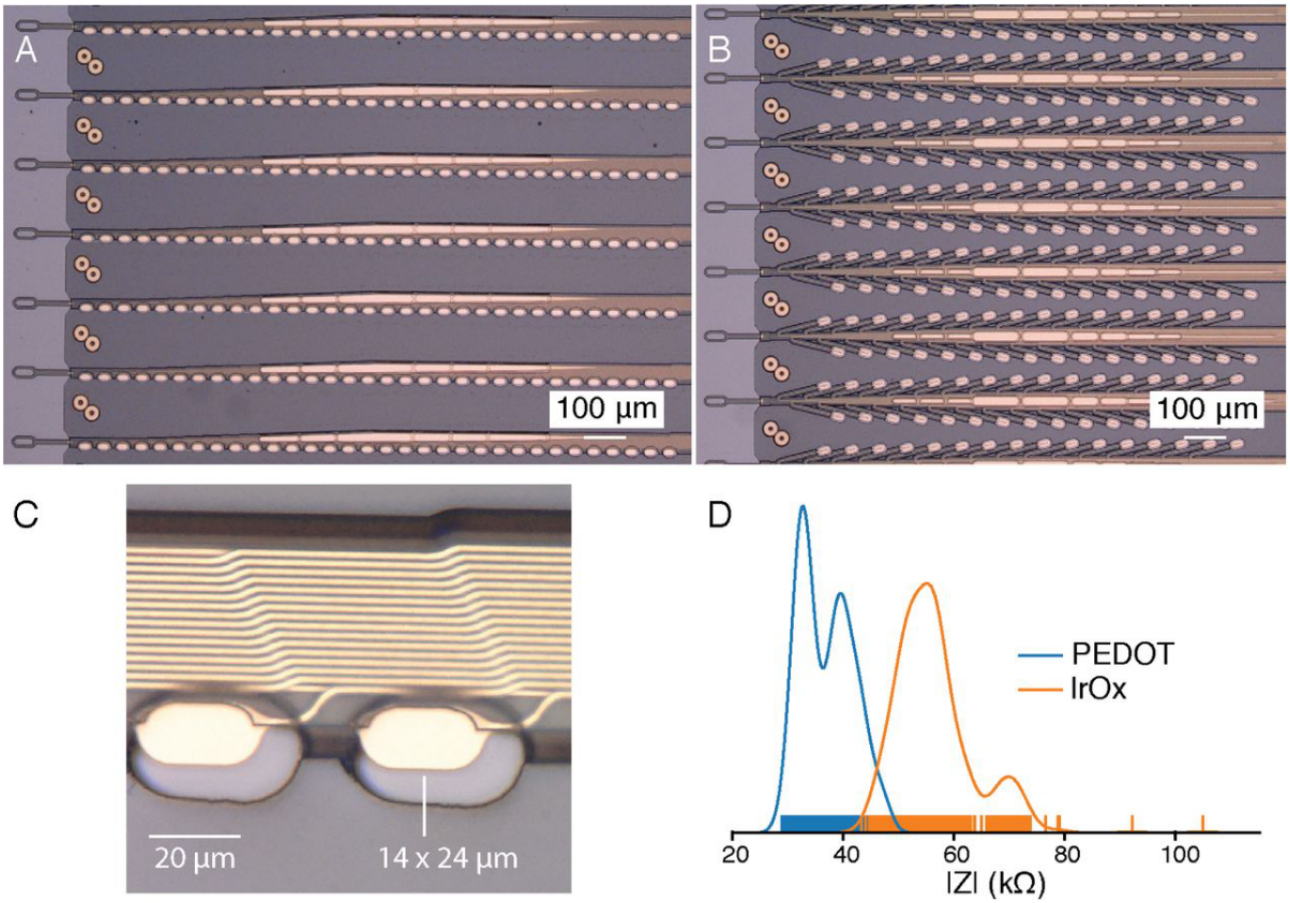
We have developed a custom process to fabricate minimally displacive neural probes that employ a variety of biocompatible thin film materials. The main substrate and dielectric used in these probes is polyimide, which encapsulates a gold thin film trace. Each thin film array is composed of a "thread" area that features electrode contacts and traces and a "sensor" area where the thin film interfaces with custom chips that enable signal amplification and acquisition. A wafer-level microfabrication process enables high-throughput manufacturing of these devices. Ten thin film devices are patterned on a wafer, each with 3072 electrode contacts.

Each array has 48 or 96 threads, each of which contain 32 independent electrodes. Integrated chips are bonded to the contacts on the sensor area of the thin film using a flip-chip bonding process. One goal of this approach is to maintain a small thread cross-sectional area to minimize tissue displacement in the brain. To achieve this, while keeping the channel count high, stepper lithography and other microfabrication techniques are used to form the metal film at submicron resolution.

We have designed and manufactured over 20 different thread and electrode types into our arrays; two example designs are shown in Figure 1A and B. Probes are designed either with the reference electrodes on separate threads or on the same threads as the recording electrodes (referred to as "on-probe references"). We have fabricated threads ranging from 5 µm to 50 µm in width that incorporate recording sites of several geometries (Figure 1). Thread thickness is nominally 4-6 µm, which includes up to three layers of insulation and two layers of conductor. Typical thread length is approximately 20 mm. To manage these long, thin threads prior to insertion, parylene-c is deposited onto the threads to form a film on which the threads remain attached until the surgical robot pulls them off. Each thread ends in a 16×50 μm^2 loop to accommodate needle threading.

Since the individual gold electrode sites have small geometric surface areas (Figure 1C), we use surface modifications to lower the impedance for electrophysiology and increase the effective charge-carrying capacity of the interface (Figure 1D). Two such treatments that we have used are the electrically conductive polymer poly-ethylenedioxythiophene doped with polystyrene sulfonate (PEDOT:PSS) [29,30] and iridium oxide (IrOx) [31,32]. In benchtop testing, we have achieved impedances of 36.97 (SD 4.68) kΩ (n=257 electrodes) and 56.46 (SD 7.10) kΩ (n=588) for PEDOT:PSS and IrOx, respectively. The lower impedance of PEDOT:PSS is promising; however, the long-term stability and biocompatibility of PEDOT:PSS are less well established than those for IrOx. These techniques and processes can be improved and further extended to other types of conductive electrode materials and coatings.

Figure 1. Our novel polymer probes. (A) "Linear Edge" probes, with 32 electrode contacts spaced by 50 µm. (B) "Tree" probes with 32 electrode contacts spaced by 75 µm. (C) Increased magnification of individual electrodes for the thread design in panel A, emphasizing their small geometric surface area. (D) Distribution of electrode impedances (measured at 1 kHz) for two surface treatments: PEDOT (n=257) and IrOx (n=588). IrOx: iridium oxide; PEDOT: poly-ethylenedioxythiophene; PCB: printed circuit board.



To keep the electronics package small, a novel alignment and flip-chip bonding process was developed. Multilevel gold stud bumps are placed throughout the printed circuit board (PCB) to act as alignment guides and temporary holders for the thin film. A custom shuttle is used to handle, align, and place the thin film on the PCB such that holes in the thin film slide around the stud bumps. The thin film is secured into place by applying force to the gold stud bumps, which flattens them into rivets. Next, the integrated chips are bonded directly to both contacts on the sensor area of the thin film and pads on the PCB by using standard flip-chip bonding processes. A custom silicon shuttle is used to vacuum pick-up rows of 40-50 capacitors and bond a total of 192 capacitors onto the PCB. This alignment and bonding process was key to creating a package containing 3072 channels in a 23×18.5 mm^2 footprint.

Robot

Thin-film polymers have previously been used for electrode probes [21], but their low bending stiffness complicates insertions. Neuralink has developed a robotic insertion approach for inserting flexible probes [28], allowing rapid and reliable insertion of large numbers of polymer probes targeted to avoid vasculature and record from dispersed brain regions. The robot's insertion head is mounted on a globally accurate, 400×400×150 mm travel, 10-µm three-axis stage and holds a small, quick-swappable "needle-pincher" assembly (Figures 2 and 3A).

The needle is milled from 40-µm diameter tungsten-rhenium wire-stock electrochemically etched to 24-µm diameter along the inserted length (Figure 2A). The tip of the needle is designed both to hook onto insertion loops—for transporting and inserting individual threads—and to penetrate the meninges and brain tissue. The needle is driven by a linear motor, allowing variable insertion speeds and rapid retraction acceleration (up to 30,000 mm/s^2) to encourage separation of the probe from the needle. The pincher is a 50-µm tungsten wire bent at the tip and driven both axially and rotationally (Figure 2B). It serves as a support for probes during transport and as a guide to ensure that threads are inserted along the needle path. Figure 3 shows a sequence of photographs of the insertion process into an agarose brain proxy.

Figure 2. Needle pincher cartridge compared with a penny for scale. (A) Needle. (B) Pincher. (C) Cartridge.

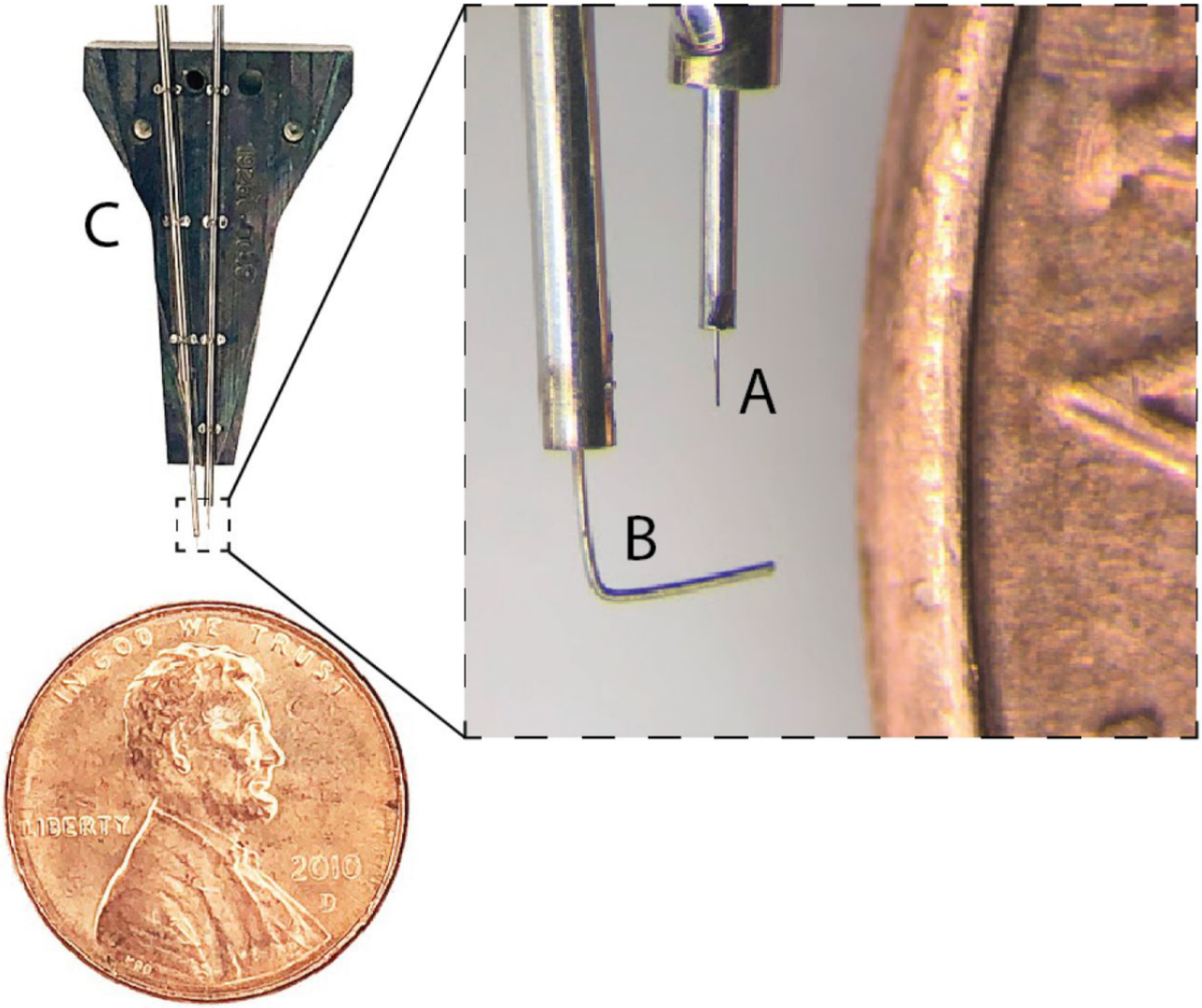
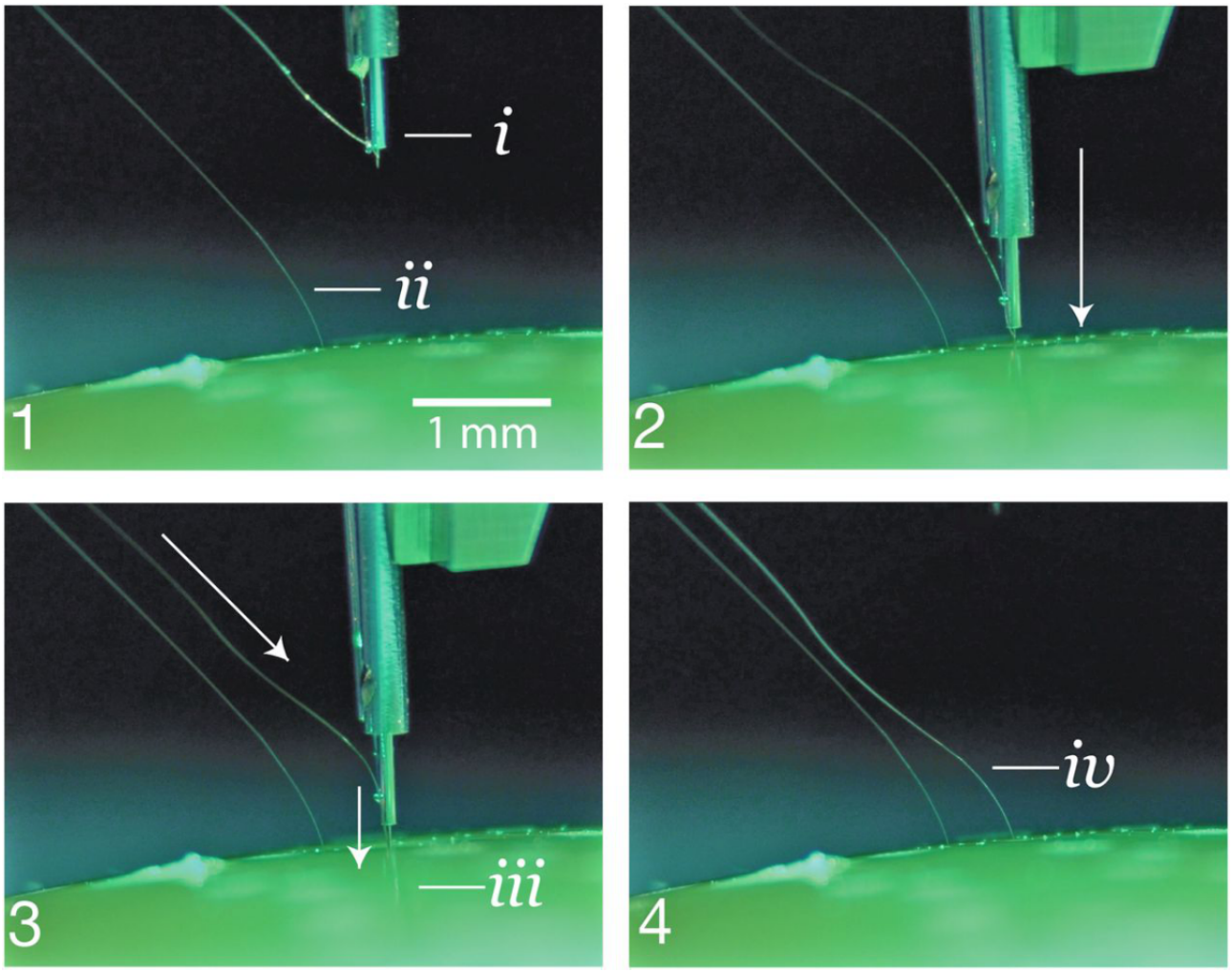
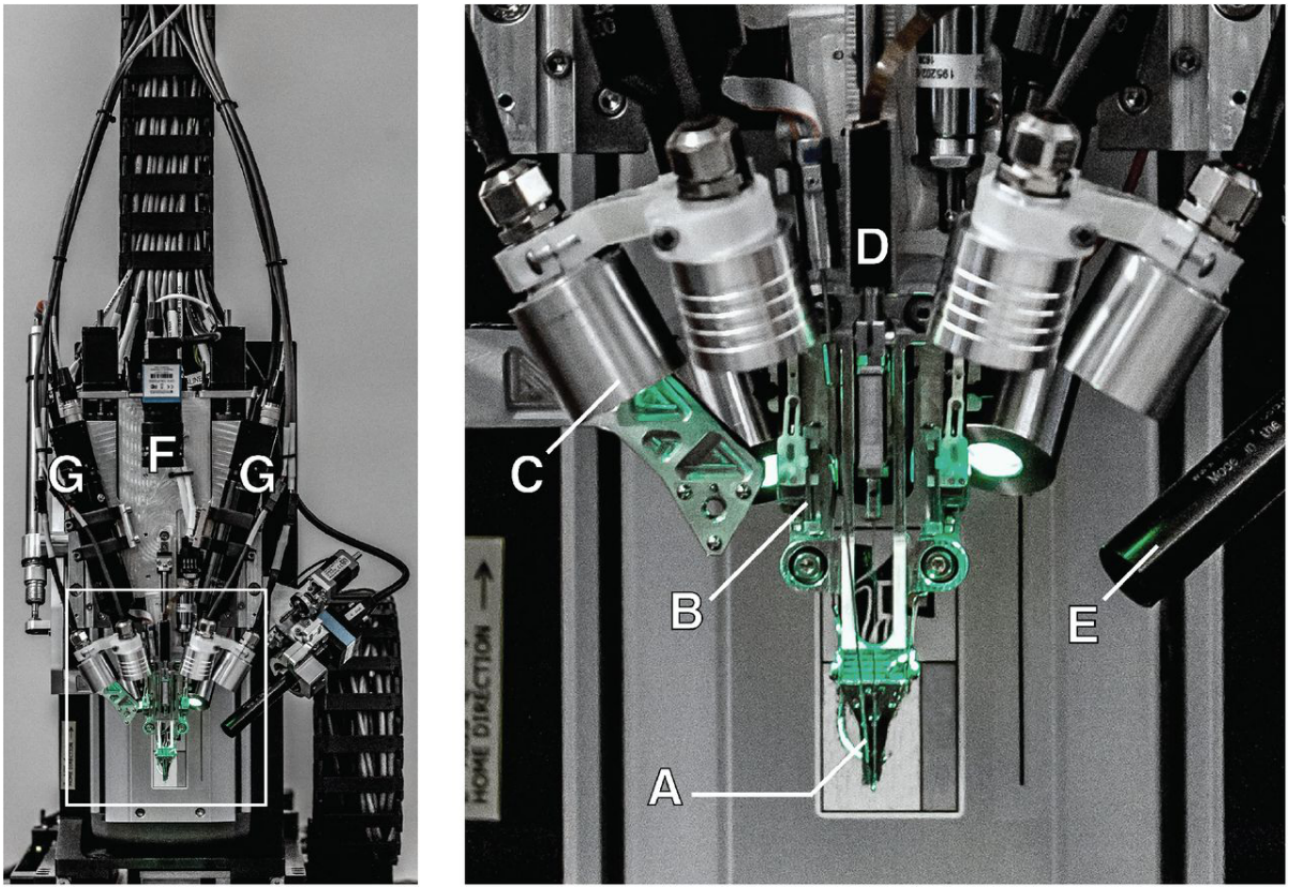


Figure 3. Insertion process into an agarose brain proxy. (1) The inserter approaches the brain proxy with a thread. (i) needle and cannula. (ii) Previously inserted thread. (2) Inserter touches down on the brain proxy surface. (3) Needle penetrates tissue proxy, advancing the thread to the desired depth. (iii) Inserting thread. (4) Inserter pulls away, leaving the thread behind in the tissue proxy. (iv) Inserted thread.



The inserter head also holds an imaging stack (Figure 4E-G) used for guiding the needle into the thread loop, insertion targeting, live insertion viewing, and insertion verification. In addition, the inserter head contains six independent light modules, each capable of independently illuminating with 405 nm, 525 nm, and 650 nm or white light (Figure 4C). The 405-nm illumination excites fluorescence from polyimide and allows the optical stack and computer vision to reliably localize the $16 \times 50\ \mu m^2$ thread loop and execute submicron visual servoing to guide, while illuminated by 650 nm light, the needle through it. Stereoscopic cameras, software-based monocular extended depth-of-field calculations, and illumination with 525 nm light allow for precise estimation of the location of the cortical surface.

Figure 4. The robotic electrode inserter; enlarged view of the inserter-head shown in the inset. (A) Loaded needle pincher cartridge. (B) Low-force contact brain position sensor. (C) Light modules with multiple independent wavelengths. (D) Needle motor. (E) One of four cameras focused on the needle during insertion. (F) Camera with wide angle view of the surgical field. (G) Stereoscopic cameras.

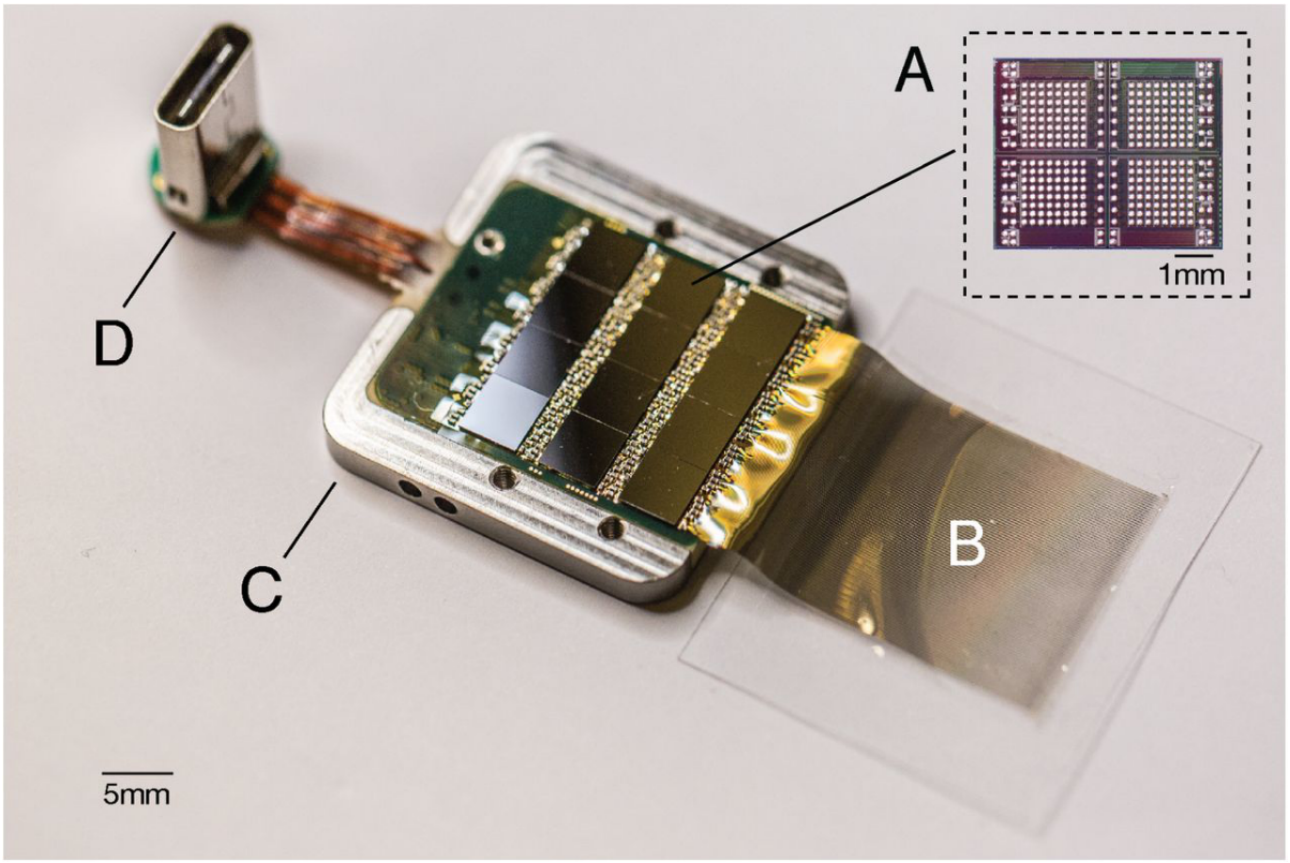


The robot registers insertion sites to a common coordinate frame with landmarks on the skull, which, when combined with depth tracking, enables precise targeting of anatomically defined brain structures. An integrated custom software suite allows preselection of all insertion sites, enabling planning of insertion paths optimized to minimize tangling and strain on the threads. The planning feature highlights the ability to avoid vasculature during insertions, one of the key advantages of inserting electrodes individually. This is particularly important, since damage to the blood-brain barrier is thought to play a key role in the brain's inflammatory response to foreign objects [33].

The robot features an autoinsertion mode, which can insert up to six threads (192 electrodes) per minute. Although the entire insertion procedure can be automated, the surgeon retains full control, and, if desired, can make manual microadjustments to the thread position before each insertion into the cortex. The neurosurgical robot is compatible with sterile shrouding and has features to facilitate successful and rapid insertions such as automatic sterile ultrasonic cleaning of the needle. The needle pincher cartridge (Figure 2C) is the portion of the inserter head that makes direct contact with brain tissue and is a consumable that can be replaced midsurgery in under a minute.

With this system, we have demonstrated an average of 87.1% (SD 12.6%) insertion success rate over 19 surgeries. In this study, precise manual adjustments were made to avoid microvasculature on the cortical surface, slowing total insertion time from the fastest possible time. Even with these adjustments, the total insertion time for this study averaged approximately 45 min for an approximate insertion rate of 29.6 electrodes per minute (Figure 5). Insertions were made in a 4×7 mm^2 bilateral craniotomy with >300 μm spacing between threads to maximize cortical coverage. This demonstrates that robotic insertion of thin polymer electrodes is an efficient and scalable approach for recording from large numbers of neurons in anatomically defined brain regions.

Figure 5. A packaged sensor device. (A) Individual neural processing application-specific integrated circuit capable of processing 256 channels of data. This particular packaged device contains 12 of these chips for a total of 3072 channels. (B) Polymer threads on parylene-c substrate. (C) Titanium enclosure (lid removed). (D) Digital USB-C connector for power and data.



Electronics

Chronic recording from thousands of electrode sites presents significant electronics and packaging challenges. The density of recording channels necessitates placing the signal amplification and digitization stack within the array assembly; otherwise, the cable and connector requirements would be prohibitive. This recording stack must amplify small neural signals (<10 μV_{RMS}) while rejecting out-of-band noise, sample and digitize the amplified signals, and stream out the results for real-time processing—all using minimal power and size.

The electronics are built around our custom Neuralink application-specific integrated circuit (ASIC), which consists of 256 individually programmable amplifiers (“analog pixels”), on-chip analog-to-digital converters (ADCs), and peripheral control circuitry for serializing the digitized outputs. The analog pixel is highly configurable: The gains and filter properties can be calibrated to account for variability in signal quality due to process variations and the electrophysiological environment. The on-chip ADC samples at 19.3 kHz with 10-bit resolution. Each analog pixel consumes 5.2 µW, and the whole ASIC consumes approximately 6 mW, including the clock drivers. Performance of the Neuralink ASIC is summarized in Table 1, and a photograph of the fabricated device is shown in Figure 6A.

The Neuralink ASIC forms the core of a modular recording platform that allows for easy replacement of constitutive parts for research and development purposes (Figure 6). In the systems discussed here, a number of ASICs are integrated into a standard PCB using flip-chip integration. Each system consists of a field-programmable gate array; real-time temperature, accelerometer, and magnetometer sensors; and a single USB-C connector for full-bandwidth data transfer. The systems are packaged in titanium cases that are coated with parylene-c, which serves as a moisture barrier to prevent fluid ingress and prolong functional lifetime.

We describe two such configurations that we have built—a 1536-channel recording system (“System A”) and a 3072-channel recording system (“System B”)—summarized in Table 2. System A employs the current-generation Neuralink ASIC, while System B uses an earlier revision with comparable functionality but poorer performance specifications.

Table 1. Neuralink application-specific integrated circuit.

Variable	Value
Number of channels	256
Gain, dB	42.9-59.4
Bandwidth, kHz	3-27
Input-referred noise (3 Hz-10 kHz), μV_{RMS}	5.9
Maximum differential input range, mV_{PP}	7.2
Analog-to-digital converter resolution, bit	10
Analog pixel power, µW	5.2

Figure 6. Thread implantation and packaging. (A) An example perioperative image showing the cortical surface with implanted threads and minimal bleeding. (B) Packaged sensor device ("System B") chronically implanted in a rat.

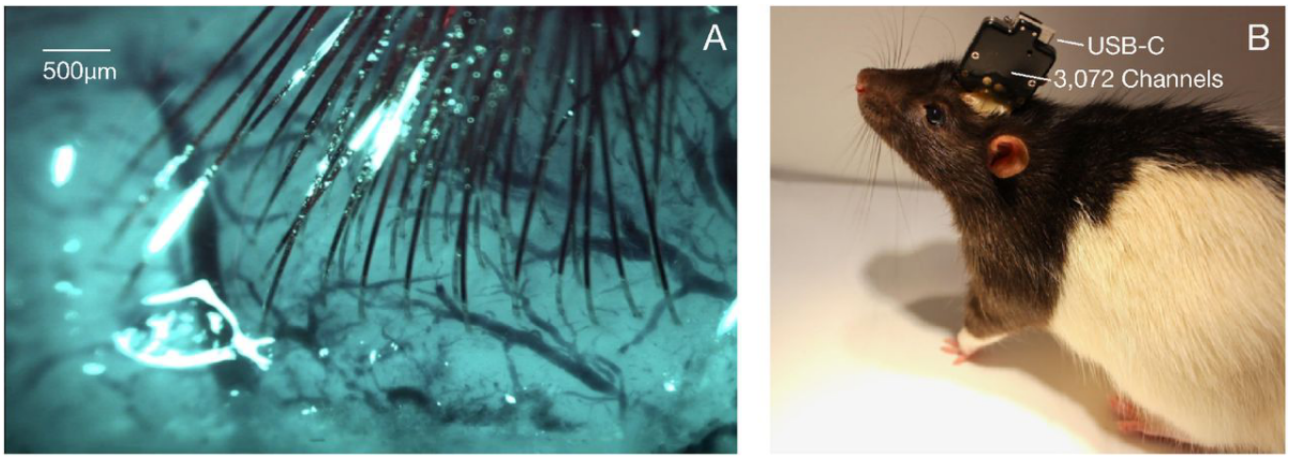


Table 2. Two recording system configurations.

Variable	Value	
	System A	System B
Number of channels	1536	3072
Sampling rate, kHz	19.3	18.6
Total system power consumption, mW	550	750
Total system size, mm^3	24.5×20×1.65	23×18.5×2
Implant weight, g	11	15

System B was designed to maximize channel density and is used for applications that demand extremely high channel count. In contrast, System A was designed to facilitate faster and more reliable manufacturing; it can be built five times faster than System B with better yields.

An Ethernet-connected base station converts the data streams from these systems into multicast 10 GB Ethernet user datagram protocol packets, allowing downstream users to process the data in a variety of ways, for example, visualizing the data in real time [34] or writing the data to disk. Each base station can connect to up to three implants simultaneously. These devices are further supported by a software ecosystem that allows for plug and play usability with zero configuration: Neural data begin streaming automatically when a cable is connected.

Electrophysiology

We have implanted both Systems A and B in male Long-Evans rats, as described in the section Robot. All animal procedures were performed in accordance with the National Research Council's *Guide for the Care and Use of Laboratory Animals* and were approved by the Neuralink Institutional Animal Care and Use Committee. Electrophysiological recordings were taken as the animals freely explored an arena equipped with a commutated cable that permitted unrestricted movement. System A can record 1344 of 1536 channels simultaneously; the exact channel configuration can be arbitrarily specified at the time of recording; System B can record from all 3072 channels simultaneously. Digitized broadband signals were processed in real time to identify action potentials (spikes) using an online detection algorithm.

Spike detection requirements for real-time brain-machine interface are different from most conventional neurophysiology requirements. While most electrophysiologists spike-sort data offline and spend significant effort to reject false-positive spike events, brain-machine interface events must be detected in real time and spike detection parameters must maximize decoding efficacy. Using our custom online spike-detection software, we found that a permissive filter that allows an estimated false-positive rate of approximately 0.2 Hz performs better than setting stringent thresholds that may reject real spikes (data not shown).

Given these considerations, we set a threshold of >0.35 Hz to quantify the number of electrodes that recorded spiking units. Since we typically do not spike sort our data, we do not report multiple units per channel. Brain-machine interface decoders commonly operate without spike sorting with minimal loss of performance [35,36]. Moreover, recent results show that spike sorting is not necessary to accurately estimate neural population dynamics [37].

Data from a recent experiment using System A are shown in Figures 7 and 8. In this experiment, 40 of 44 attempted insertions were successful (90%) for a total of 1280 implanted electrodes, of which 1020 were recorded simultaneously. The broadband signals recorded from a representative thread show both local field and spiking activity (Figure 7). A sample output of the spike detection pipeline is shown in raster form in Figure 8. In this example, two overlapping recording configurations were used to record from all 1280 implanted channels. On this array, our spiking yield was 43.4% of the channels, with many spikes appearing on multiple neighboring channels, as has been observed in other high-density probes [16,17,21]. On other System A arrays, we observed a spiking yield of 45.60% (SD 0.03%) across 19 surgeries, with a maximum spiking yield of 70%.

Figure 7. The broadband signals recorded from a representative thread. Left: Broadband neural signals (unfiltered) simultaneously acquired from a single thread (32 channels) implanted in rat cerebral cortex. Each channel (row) corresponds to an electrode site on the thread (schematic at left; sites spaced by 50 µm). Spikes and local field potentials are readily apparent. Right: Putative waveforms (unsorted); numbers indicate channel location on thread. Mean waveform is shown in black.

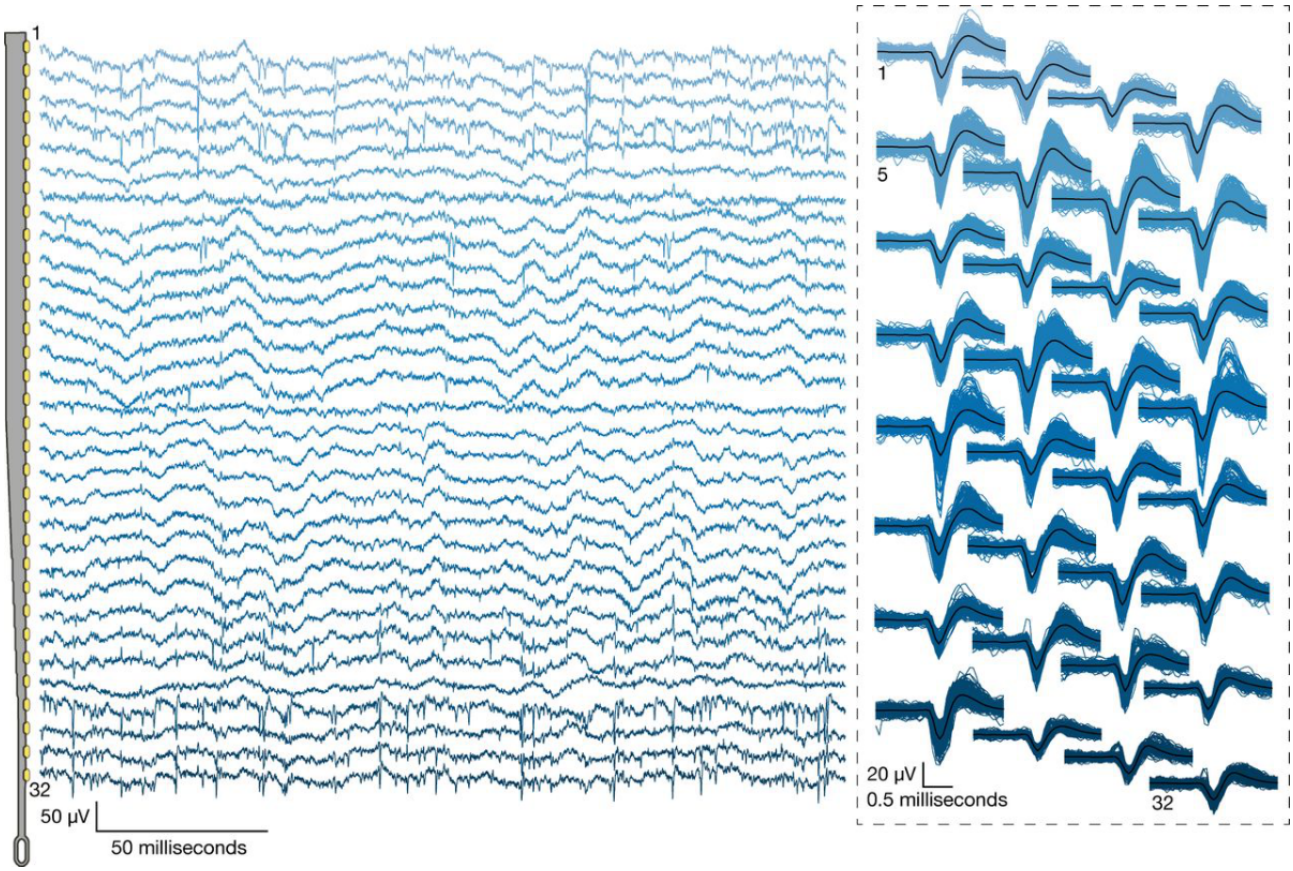
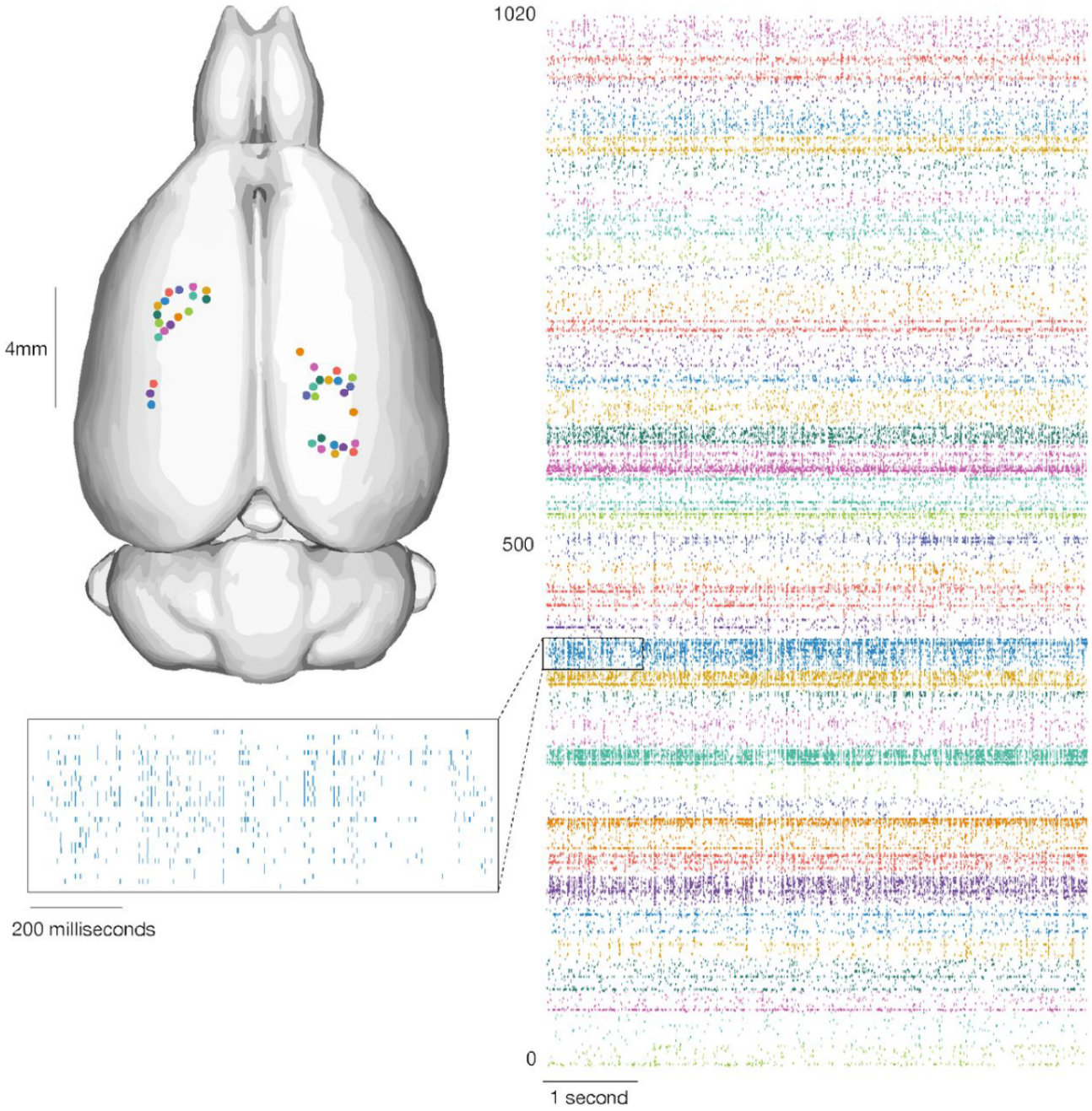


Figure 8. Our devices allow the recording of widespread neural activity distributed across multiple brain regions and cortical layers. Left: Thread insertion sites (colored circles) are indicated on the rendered rodent brain [38]. Right: Raster of 1020 simultaneously recorded channels, sorted per thread (color corresponds to insertion site). Inset: Enlarged raster of spikes from a single thread. This thread corresponds to the one shown in Figure 7.



Discussion

We have described a brain-machine interface with a high-channel count and single-spike resolution. It is based on flexible polymer probes, a robotic insertion system, and custom low-power electronics. This system serves two main purposes: It is a research platform for use in rodents and serves as a prototype for future human clinical implants. The ability to quickly iterate designs and testing in rodents allows for the rapid refinement of devices, manufacturing processes, and software. Because it is a research platform, the system uses a wired connection to maximize the bandwidth for raw data streaming. This is important for performance assessments and crucial for the development of signal processing and decoding algorithms. In contrast, the clinical devices that derive from this platform will be fully implantable, which requires hermetic packaging, and have on-board signal compression, reduced power consumption, wireless power transmission, and data telemetry through the skin without percutaneous leads.

Modulating neural activity will be an important part of next-generation clinical brain-machine interfaces [39], for example, to provide a sense of touch or proprioception to neuroprosthetic movement control [40,41]. Therefore, we designed the Neuralink ASIC to be capable of electrical stimulation on every channel, although we have not demonstrated these capabilities here.

This brain-machine interface system has several advantages over previous approaches. The size and composition of the thin-film probes are a better match for the material properties

of brain tissue than commonly used silicon probes and may therefore exhibit enhanced biocompatibility [28,21]. In addition, the ability to choose where our probes are inserted, including into the subcortical structures, allows us to create custom array geometries for targeting specific brain regions while avoiding vasculature. This feature is significant for creating a high-performance brain-machine interface, as the distribution of electrodes can be customized depending on the task requirements. Lastly, the miniaturization and design of the Neuralink ASIC affords great flexibility in system design and supports very high channel counts within practical size and power constraints.

In principle, our approach to brain-machine interfaces is highly extensible and scalable. Here, we report simultaneous broadband recording from over 3000 inserted electrodes in a freely moving rat. In a larger brain, multiple devices with this architecture could be readily implanted, and we could therefore interface with many more neurons without extensive re-engineering. Further development of surgical robotics could allow us to accomplish this without dramatically increasing surgery time.

Although significant technological challenges must be addressed before a high-bandwidth device is suitable for clinical application, with such a device, it is plausible to imagine that a patient with spinal cord injury could dexterously control a digital mouse and keyboard. When combined with rapidly improving spinal stimulation techniques [42], in the future, this approach could conceivably restore motor function. High-bandwidth neural interfaces should enable a variety of novel therapeutic possibilities.

Acknowledgments

We would like to acknowledge the contributions of Lawrence Livermore National Laboratory (LLNL), Berkeley Marvell Nanofabrication Laboratory, Berkeley Wireless Research Center (BWRC), Stanford Nanofabrication Facility, and former and current Neuralink employees in the work described here. Thread manufacturing and high-density package assembly were performed in collaboration with LLNL (CRADA No. TC02267).

Conflicts of Interest

Authors are affiliated with Neuralink.

Multimedia Appendix 1

Video 1: A series of six insertions by the neurosurgical robot into an agarose brain proxy. Thread-capture by the needle occurs off-frame. The changes in background color are caused by illumination with different frequencies of light at different stages of the threading and insertion process. One thread was inserted before the start of the video.

[MP4 File (MP4 Video), 17150 KB - jmir_v21i10e16194_app1.mp4]

Multimedia Appendix 2

Video 2: A 3D rendered view of thread arrangement (same data presented in Figure 8). Thread insertion is visualized in the same order as in the actual surgery, but time has been compressed for presentation. Thread size and insertion depth are representative. The stereotaxic coordinates of each insertion are represented on the dataset provided by Calabrese and coworkers [35].

[MP4 File (MP4 Video), 1617 KB - jmir_v21i10e16194_app2.mp4]

References

1. Hochberg LR, Serruya MD, Friehs GM, Mukand JA, Saleh M, Caplan AH, et al. Neuronal ensemble control of prosthetic devices by a human with tetraplegia. Nature 2006 Jul 13;442(7099):164-171. [doi: 10.1038/nature04970] [Medline: 16838014]
2. Wang W, Collinger JL, Degenhart AD, Tyler-Kabara EC, Schwartz AB, Moran DW, et al. An electrocorticographic brain interface in an individual with tetraplegia. PLoS One 2013;8(2):e55344 [FREE Full text] [doi: 10.1371/journal.pone.0055344] [Medline: 23405137]
3. Aflalo T, Kellis S, Klaes C, Lee B, Shi Y, Pejsa K, et al. Neurophysiology. Decoding motor imagery from the posterior parietal cortex of a tetraplegic human. Science 2015 May 22;348(6237):906-910 [FREE Full text] [doi: 10.1126/science.aaa5417] [Medline: 25999506]
4. Hochberg LR, Bacher D, Jarosiewicz B, Masse NY, Simeral JD, Vogel J, et al. Reach and grasp by people with tetraplegia using a neurally controlled robotic arm. Nature 2012 May 16;485(7398):372-375 [FREE Full text] [doi: 10.1038/nature11076] [Medline: 22596161]
5. Collinger JL, Wodlinger B, Downey JE, Wang W, Tyler-Kabara EC, Weber DJ, et al. High-performance neuroprosthetic control by an individual with tetraplegia. Lancet 2013 Feb 16;381(9866):557-564 [FREE Full text] [doi: 10.1016/S0140-6736(12)61816-9] [Medline: 23253623]
6. Anumanchipalli GK, Chartier J, Chang EF. Speech synthesis from neural decoding of spoken sentences. Nature 2019 Apr;568(7753):493-498. [doi: 10.1038/s41586-019-1119-1] [Medline: 31019317]

7. Buzsáki G, Anastassiou CA, Koch C. The origin of extracellular fields and currents--EEG, ECoG, LFP and spikes. Nat Rev Neurosci 2012 May 18;13(6):407-420 [FREE Full text] [doi: 10.1038/nrn3241] [Medline: 22595786]
8. Pesaran B, Vinck M, Einevoll GT, Sirota A, Fries P, Siegel M, et al. Investigating large-scale brain dynamics using field potential recordings: analysis and interpretation. Nat Neurosci 2018 Jul;21(7):903-919. [doi: 10.1038/s41593-018-0171-8] [Medline: 29942039]
9. Kaiju T, Doi K, Yokota M, Watanabe K, Inoue M, Ando H, et al. High spatiotemporal resolution ECoG recording of somatosensory evoked potentials with flexible micro-electrode arrays. Front Neural Circuits 2017;11:20 [FREE Full text] [doi: 10.3389/fncir.2017.00020] [Medline: 28442997]
10. Yuste R. From the neuron doctrine to neural networks. Nat Rev Neurosci 2015 Aug;16(8):487-497. [doi: 10.1038/nrn3962] [Medline: 26152865]
11. Hong G, Lieber CM. Novel electrode technologies for neural recordings. Nat Rev Neurosci 2019 Jun;20(6):330-345 [FREE Full text] [doi: 10.1038/s41583-019-0140-6] [Medline: 30833706]
12. Maynard EM, Nordhausen CT, Normann RA. The Utah intracortical Electrode Array: a recording structure for potential brain-computer interfaces. Electroencephalogr Clin Neurophysiol 1997 Mar;102(3):228-239. [doi: 10.1016/s0013-4694(96)95176-0] [Medline: 9129578]
13. Nicolelis MAL, Dimitrov D, Carmena JM, Crist R, Lehew G, Kralik JD, et al. Chronic, multisite, multielectrode recordings in macaque monkeys. Proceedings of the National Academy of Sciences 2003 Sep 05;100(19):11041-11046 [FREE Full text] [doi: 10.1073/pnas.1934665100]
14. Wise K, Sodagar A, Yao Y, Gulari M, Perlin G, Najafi K. Microelectrodes, microelectronics, and implantable neural microsystems. Proceedings of the IEEE 2008 Jul;96(7):1184-1202. [doi: 10.1109/jproc.2008.922564]
15. Dotson NM, Hoffman SJ, Goodell B, Gray CM. A large-scale semi-chronic microdrive recording system for non-human primates. Neuron 2017 Nov 15;96(4):769-782.e2 [FREE Full text] [doi: 10.1016/j.neuron.2017.09.050] [Medline: 29107523]
16. Jun JJ, Steinmetz NA, Siegle JH, Denman DJ, Bauza M, Barbarits B, et al. Fully integrated silicon probes for high-density recording of neural activity. Nature 2017 Nov 08;551(7679):232-236 [FREE Full text] [doi: 10.1038/nature24636] [Medline: 29120427]
17. Angotzi GN, Boi F, Lecomte A, Miele E, Malerba M, Zucca S, et al. SiNAPS: An implantable active pixel sensor CMOS-probe for simultaneous large-scale neural recordings. Biosens Bioelectron 2019 Feb 01;126:355-364 [FREE Full text] [doi: 10.1016/j.bios.2018.10.032] [Medline: 30466053]
18. Deku F, Cohen Y, Joshi-Imre A, Kanneganti A, Gardner TJ, Cogan SF. Amorphous silicon carbide ultramicroelectrode arrays for neural stimulation and recording. J Neural Eng 2018 Feb;15(1):016007 [FREE Full text] [doi: 10.1088/1741-2552/aa8f8b] [Medline: 28952963]
19. Lecomte A, Descamps E, Bergaud C. A review on mechanical considerations for chronically-implanted neural probes. J Neural Eng 2018 Jun;15(3):031001. [doi: 10.1088/1741-2552/aa8b4f] [Medline: 28885187]
20. Khodagholy D, Gelinas JN, Thesen T, Doyle W, Devinsky O, Malliaras GG, et al. NeuroGrid: recording action potentials from the surface of the brain. Nat Neurosci 2014 Dec 22;18(2):310-315. [doi: 10.1038/nn.3905]
21. Chung JE, Joo HR, Fan JL, Liu DF, Barnett AH, Chen S, et al. High-density, long-lasting, and multi-region electrophysiological recordings using polymer electrode arrays. Neuron 2019 Jan 02;101(1):21-31.e5. [doi: 10.1016/j.neuron.2018.11.002] [Medline: 30502044]
22. Chorover SL, Deluca A. A sweet new multiple electrode for chronic single unit recording in moving animals. Physiology & Behavior 1972 Oct;9(4):671-674. [doi: 10.1016/0031-9384(72)90030-3]
23. Liu J, Fu T, Cheng Z, Hong G, Zhou T, Jin L, et al. Syringe-injectable electronics. Nature Nanotech 2015 Jun 8;10(7):629-636. [doi: 10.1038/nnano.2015.115]
24. Fu T, Hong G, Zhou T, Schuhmann TG, Viveros RD, Lieber CM. Stable long-term chronic brain mapping at the single-neuron level. Nat Methods 2016 Oct;13(10):875-882. [doi: 10.1038/nmeth.3969] [Medline: 27571550]
25. Vitale F, Vercosa DG, Rodriguez AV, Pamulapati SS, Seibt F, Lewis E, et al. Fluidic microactuation of flexible electrodes for neural recording. Nano Lett 2018 Jan 10;18(1):326-335 [FREE Full text] [doi: 10.1021/acs.nanolett.7b04184] [Medline: 29220192]
26. Luan L, Wei X, Zhao Z, Siegel JJ, Potnis O, Tuppen CA, et al. Ultraflexible nanoelectronic probes form reliable, glial scar-free neural integration. Sci Adv 2017 Feb;3(2):e1601966 [FREE Full text] [doi: 10.1126/sciadv.1601966] [Medline: 28246640]
27. Marc D. BioRxiv. 2018. NeuroRoots, a bio-inspired, seamless brain machine interface device for long-term recording URL: https://www.biorxiv.org/content/biorxiv/early/2018/11/04/460949.full.pdf [accessed 2019-09-11]
28. Timothy LH. BioRxiv. 2019. The “sewing machine” for minimally invasive neural recording URL: https://www.biorxiv.org/content/biorxiv/early/2019/03/14/578542.full.pdf [accessed 2019-09-11]
29. Ludwig KA, Uram JD, Yang J, Martin DC, Kipke DR. Chronic neural recordings using silicon microelectrode arrays electrochemically deposited with a poly(3,4-ethylenedioxythiophene) (PEDOT) film. J Neural Eng 2006 Mar;3(1):59-70. [doi: 10.1088/1741-2560/3/1/007] [Medline: 16510943]

30. Wilks SJ, Richardson-Burns SM, Hendricks JL, Martin DC, Otto KJ. Poly(3,4-ethylenedioxythiophene) as a micro-neural interface material for electrostimulation. Front Neuroeng 2009;2:7 [FREE Full text] [doi: 10.3389/neuro.16.007.2009] [Medline: 19543541]
31. Klein JD, Clauson SL, Cogan SF. Morphology and charge capacity of sputtered iridium oxide films. Journal of Vacuum Science & Technology A: Vacuum, Surfaces, and Films 1989 Sep;7(5):3043-3047. [doi: 10.1116/1.576313]
32. Cogan SF, Plante TD, Ehrlich J. Sputtered Iridium Oxide Films (SIROFs) for Low-Impedance Neural Stimulation and Recording Electrodes. 2004 Presented at: The 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society; 2004; San Francisco, USA. [doi: 10.1109/iembs.2004.1404158]
33. Saxena T, Karumbaiah L, Gaupp EA, Patkar R, Patil K, Betancur M, et al. The impact of chronic blood-brain barrier breach on intracortical electrode function. Biomaterials 2013 Jul;34(20):4703-4713. [doi: 10.1016/j.biomaterials.2013.03.007] [Medline: 23562053]
34. Siegle JH, López AC, Patel YA, Abramov K, Ohayon S, Voigts J. Open Ephys: an open-source, plugin-based platform for multichannel electrophysiology. J Neural Eng 2017 Aug;14(4):045003. [doi: 10.1088/1741-2552/aa5eea] [Medline: 28169219]
35. Todorova S, Sadtler P, Batista A, Chase S, Ventura V. To sort or not to sort: the impact of spike-sorting on neural decoding performance. J Neural Eng 2014 Oct;11(5):056005 [FREE Full text] [doi: 10.1088/1741-2560/11/5/056005] [Medline: 25082508]
36. Christie BP, Tat DM, Irwin ZT, Gilja V, Nuyujukian P, Foster JD, et al. Comparison of spike sorting and thresholding of voltage waveforms for intracortical brain–machine interface performance. J Neural Eng 2014 Dec 11;12(1):016009. [doi: 10.1088/1741-2560/12/1/016009]
37. Trautmann EM, Stavisky SD, Lahiri S, Ames KC, Kaufman MT, O'Shea DJ, et al. Accurate estimation of neural population dynamics without spike sorting. Neuron 2019 Jul 17;103(2):292-308.e4. [doi: 10.1016/j.neuron.2019.05.003] [Medline: 31171448]
38. Calabrese E, Badea A, Watson C, Johnson GA. A quantitative magnetic resonance histology atlas of postnatal rat brain development with regional estimates of growth and variability. Neuroimage 2013 May 01;71:196-206 [FREE Full text] [doi: 10.1016/j.neuroimage.2013.01.017] [Medline: 23353030]
39. Zhou A, Santacruz SR, Johnson BC, Alexandrov G, Moin A, Burghardt FL, et al. A wireless and artefact-free 128-channel neuromodulation device for closed-loop stimulation and recording in non-human primates. Nat Biomed Eng 2018 Dec 31;3(1):15-26. [doi: 10.1038/s41551-018-0323-x]
40. O'Doherty JE, Lebedev MA, Ifft PJ, Zhuang KZ, Shokur S, Bleuler H, et al. Active tactile exploration using a brain–machine–brain interface. Nature 2011 Oct 5;479(7372):228-231. [doi: 10.1038/nature10489]
41. Sharlene NF. BioRxiv. 2019. Restored tactile sensation improves neuroprosthetic arm control? URL: https://www.biorxiv.org/content/biorxiv/early/2019/05/31/653428.full.pdf [accessed 2019-09-11]
42. Wagner FB, Mignardot J, Le Goff-Mignardot CG, Demesmaeker R, Komi S, Capogrosso M, et al. Targeted neurotechnology restores walking in humans with spinal cord injury. Nature 2018 Oct 31;563(7729):65-71. [doi: 10.1038/s41586-018-0649-2]

Abbreviations

ADC: analog-to-digital converters
ASIC: application-specific integrated circuit
IrOx: iridium oxide
PCB: printed circuit board
PEDOT:PSS: poly-ethylenedioxythiophene doped with polystyrene sulfonate

Edited by G Eysenbach; submitted 09.09.19; peer-reviewed by G Valle; comments to author 26.09.19; accepted 14.10.19; published 31.10.19.

Please cite as:
Musk E, Neuralink
An Integrated Brain-Machine Interface Platform With Thousands of Channels
J Med Internet Res 2019;21(10):e16194
URL: http://www.jmir.org/2019/10/e16194/
doi:10.2196/16194
PMID:31642810

Commentary

The Ethical and Responsible Development and Application of Advanced Brain Machine Interfaces

Andrew David Maynard[1*], BSc, PhD; Marissa Scragg[1*], BA

School for the Future of Innovation in Society, Arizona State University, Tempe, AZ, United States

[*]all authors contributed equally

Corresponding Author:
Andrew David Maynard, BSc, PhD
School for the Future of Innovation in Society
Arizona State University
PO Box 875603
Tempe, AZ, 85287-5603
United States
Phone: 1 480 727 8831
Fax: 1 480 727 8791
Email: andrew.maynard@asu.edu

Related Article:

Comment on: http://jmir.org/2019/10/e16194/

Abstract

Advanced brain machine interfaces provide potentially transformative approaches to treating neurological conditions and enhancing the performance of users. Yet, as technological capabilities continue to progress in leaps and bounds, there is a possibility that these capabilities outstrip our collective understanding of how to ensure brain machine interfaces are developed and used ethically and responsibly. In this case, there is an overt danger of rapid technological developments leading to unanticipated harm through a lack of foresight including threats to privacy, autonomy, self-identity, and other areas of personal and social value which, while hard to quantify, represent substantial risks. There is also a very real likelihood of such risks undermining value creation around the technologies and the associated enterprises, as key stakeholders push back against perceived and actual threats to what they, in turn, hold to be of value. In order to successfully traverse the resulting risk landscape, researchers and developers will need to become increasingly adept at integrating a sophisticated understanding of ethical and socially responsible innovation into their enterprises. Here, we illustrate how a "risk innovation" approach may provide novel insights into mapping out this landscape and revealing potentially blindsiding risks. We show how this approach can be used to illuminate challenges and opportunities to the successful, ethical, and responsible development of advanced brain machine interfaces. In addition, we emphasize how success will ultimately depend on the willingness of innovators and others to take ethical and responsible innovation seriously and to draw on the interdisciplinary and transdisciplinary expertise that is necessary to translate good intentions into positive outcomes.

(J Med Internet Res 2019;21(10):e16321) doi:10.2196/16321

KEYWORDS

brain machine interface; ethics; neuroethics; bioethics; ethical innovation; responsible innovation; risk; risk innovation

Introduction

Invasive technologies that enable machines to directly interface with regions of the brain have been the subject of research and development for some decades. The invention of cochlear implants in 1957 was a pivotal point in brain machine interface development [1], and since then, an increasingly sophisticated array of brain machine interface technologies have emerged, from deep brain stimulation techniques [2] to the high-density microelectrode arrays such as the Utah Array [3]. With intensifying interest in neurotechnologies reflected, in part, in large-scale research initiatives such as the US Brain Initiative and the Human Brain Project in Europe, brain machine interface technologies are continuing to attract considerable attention. However, recent technical advances published by Elon Musk and colleagues [4] suggest that a watershed in brain machine interface technology is approaching, which has the potential to

not only revolutionize the treatment of neurological disorders but also transition brain machine interfaces from being firmly rooted in the domain of exclusive therapeutic devices to transformative and widely available performance-enhancing technologies.

As this technology continues to mature, there are questions associated with its ethical and responsible development and use that need to be addressed if the technology is to improve lives without causing unanticipated and potentially serious harm. Over the past several years, these have been explored extensively by ethicists, researchers, and others [5-9] and include a better understanding of the nature and acceptability of potential risks to health, behavior, and personality/sense of identity as well as issues including who benefits from the technology and whether access confers an unfair advantage on users. They also raise concerns associated with privacy, user autonomy—especially where users have limited control over implanted brain machine interfaces and the data they produce—and challenges of data and device security. On the other hand, the potential benefits of brain machine interfaces—both for therapeutic and enhancement uses—raise important ethical questions on the extent to which slowing or stopping development may disadvantage future beneficiaries of the technology. Most recently, the UK Royal Society grappled with these and similar questions in a comprehensive perspective on emerging brain machine interface technologies, concluding that "neural interface technologies will continue to raise profound ethical, political, social and commercial questions that should be addressed as soon as possible to create mechanisms to approve, regulate or control the technologies as they develop, as well as managing the impact they may have on society" [10].

Within this emerging landscape, developers of advanced brain machine interfaces are in urgent need of guidance on how to proceed appropriately if they are to ensure the value of the technology is fully realized, without overstepping ethical lines. However, few frameworks exist that explicitly aid researchers, businesses, and others in developing brain machine interface technologies that are ethical, responsible, and successful. One emerging framework that is potentially useful in revealing pathways forward around ethical and responsible development is that of "risk innovation" [11]. This is a framework we are in the process of developing and testing, which is designed to support pragmatic decisions around responsible and ethical innovation, within the often-tight constraints enterprises face as they push the bounds of what is possible. Here, we explore the applicability of a risk innovation approach to guiding the ethical and responsible development of advanced brain machine interfaces [12].

Challenges of Developing New Technologies Ethically and Responsibly

Developers of potentially transformative technologies face an increasingly complex array of challenges as they strive to balance value creation and economic success with ethical and socially responsible development. Although they often have a vision of their technology being used to improve lives or the environment, they are frequently operating under conditions of extreme uncertainty, within tight resource constraints and with somewhat limited understanding of how their technology may affect others or be perceived by them. The result is a convoluted and shifting "risk landscape" that developers need to traverse in order to be successful and one that is heavily influenced by unconventional, indistinct, and often people- and society-oriented hurdles. Traversing this landscape depends on developing and applying a sophisticated understanding of stakeholder value and values and a commitment to protecting and supporting these, where appropriate. This is a metaphorical landscape that lies between good ideas and their successful development. It is littered with potential pitfalls and often includes obstacles that are outside the immediate expertise of those attempting to traverse it. As the coupling between technology innovation and society becomes increasingly tight, this landscape shifts more rapidly than it has in the past [13].

Although quantitative evidence for how this shifting landscape impacts innovation remains elusive, it is hard to miss the challenges that technology companies such as Facebook, Google, and Amazon are now facing because of their failure to map out and plan for potentially blindsiding risks including risks that often have their roots in societal expectations and norms. This is part of a broader trend where concerns associated with issues such as privacy, justice, and autonomy have an ever-larger impact on the challenges technology companies need to navigate in order to succeed. At least some of these challenges stem from businesses focusing on *shareholder* value while failing to recognize just how quickly disregarding *stakeholder* value (including value to customers, impacted communities, and publics more generally) can shut them down. This is a trend that companies developing genetically modified products have painfully learned to account for in recent years [14]. More broadly, these trends are indicative of the ever-more complex dynamic between technology innovation and society, which is increasing the unpredictability of the risk landscape in many cases.

In part, because of the evolving complexity of this risk landscape, there is a growing recognition that long-term value creation around technology innovations depends on ethical and socially responsible development and an ability to anticipate and take early action to avoid potential issues [15]. This is seen in approaches to, for example, anticipatory governance [16] and agile governance [17], which aim to equip businesses, regulators, and others with the insights, tools, and skills to anticipate and navigate potential challenges. It is also deeply embedded in thinking around responsible innovation [18], which strives to provide developers and others with a framework within which commercial success is tied to social responsibility.

These are all approaches that are directly relevant to the development of advanced brain machine interfaces. Advanced brain machine interfaces represent technological advances which, while potentially transformative in a positive way, are likely to face nontechnical hurdles that could derail them if there is no broad and sophisticated understanding of how to develop them ethically and responsibly. However, many approaches being explored are hard to translate into concrete actions, especially within a high-speed technology sector that is on the cutting edge of redefining what is possible.

Risk Innovation

To address the broader challenges here, we are developing pragmatic approaches to innovating responsibly and successfully within the framework of "risk innovation." This is an approach to innovation that recognizes that successful and responsible development of novel technologies cannot be predicated on treating future risks the same way as past risks and assuming that established risk assessment and management approaches will continue to be applicable. Rather, it encourages innovative ways of thinking about and acting on risk, which reveal novel pathways to successful and socially responsible technology innovation [11].

At the core of risk innovation is the concept of approaching risk as a threat to value, together with metaphorically visualizing risk as a landscape to be traversed, which lies between new ideas and their translation into successful products. Here, "value" is defined by the context surrounding a new technology or product and the stakeholders and communities potentially impacted by it. In this way, value may take on conventional attributes of health, wellbeing, environmental impact, security, and economic growth. However, it may also take on equally important but often less tangible—and frequently overlooked—attributes such as dignity, identity, justice, privacy, and autonomy.

This framing of risk enables innovators and others to consider the question "risk to whom or what?" and not simply to the immediate enterprise. Within a tightly coupled world, threats to what is of value to investors, customers, and communities become threats to the enterprise or technology. For instance, a technology that threatens the privacy of key communities (or is perceived to do so) has the potential to elicit a backlash which, in turn, threatens trust in the technology and its ultimate adoption. Likewise, corporate decisions that threaten what is important to the workforce—for instance, developing military applications without transparency or consultation—has the potential to threaten the ability of a company to retain skilled employees [19].

In our work, we are developing tools and approaches that use the ideas behind risk innovation to enable innovators to better understand and navigate the risk landscape they face. Our methodology is designed to foster a risk innovation mindset. It is grounded in helping innovators first identify key areas of value to their organization, their investors, their consumers, and their communities and then identify and address "orphan risks" [20] that have the potential to threaten these areas of value. These are risks that are easy to overlook and ignore and yet have the ability to blindside development if they are not mapped out and addressed.

Within this approach, we focus on 18 specific orphan risks that span three domains covering social and ethical factors, organizations and systems, and unintended consequences of emerging technologies (Textbox 1). This is not an exhaustive list of orphan risks, and with one or two exceptions, it does not include more conventional risks for which there are established risk assessment and management tools and approaches. For example, cyber security is not explicitly listed, although less-obvious dimensions of cyber-based risks such as privacy and loss of agency are listed. Rather, these orphan risks were identified through assessment of entrepreneurial risk landscapes as being important to raising awareness around often-overlooked and hard-to-quantify risks. They are also intended to stimulate innovative approaches to identifying and navigating similar risks that do not appear on the list.

By starting with areas of value and then considering relevant orphan risks, this approach provides innovators with unique insights into potential pitfalls and steps that may be taken to navigate around them. It is an approach that is aimed at supporting successful development, while, at the same time, avoiding threats to value amongst communities that, in turn, have the ability to stymie progress. In doing so, it actively encourages ethical, responsible, and successful development. Importantly, our approach is explicitly designed to help innovators and others avoid the dangers of ignoring unfamiliar risks or paralysis by analysis as they grapple with being overwhelmed by a deluge of speculative risks.

Textbox 1. Orphan risks used within our "risk innovation" methodology that have the potential to undermine value creation in unexpected ways if not mapped out and addressed.

Organizations and systems

- Bad actors: Risks from enterprises that behave in ways that are ethically questionable or that lead to unacceptable harm.
- Geopolitics: Risks from a lack of awareness of or strategies for navigating a shifting geopolitical landscape.
- Governance and Regulation: Risks from often evolving laws, policies, and practices that govern and guide business operations.
- Organizational Values and Culture: Risks from tensions between business practices, both internal and external, and the set of values that reflect what is important to a business' founders and members.
- Reputation and Trust: Risks from a business that has only a rudimentary understanding of how their behavior and actions strengthen or weaken reputation and trust.
- Standards: Risks from a business' lack of engagement with an evolving operational framework for businesses that spans legal requirements, informal guidelines, and norms and codes.

Unintended consequences of emerging technologies

- Black Swan Events: Risks from very-low-probability but high-impact events.
- Co-opted Tech: Risks from technologies and products that are used in ways that undermine the intention of the original business or business owner.
- Heath and Environment: Risks from new technologies and the products they are associated with, behaving in sufficiently novel ways that they potentially lead to threats to human health and the environment.
- Intergenerational Impacts: Risks from technologies that have potential impacts from one generation to another.
- Loss of Agency: Risks from products or business practices that reduce the ability of organizations and individuals to have agency.
- Product Lifecycle: Risks from unintended impacts of where and how a product's materials are sourced and manufactured, how it is used, and its disposal and reuse.

Social and ethical factors

- Ethics: Risks from business practices overstepping the often-indistinct line between ethical and unethical behavior.
- Perception: Risks created from how people perceive a technology to impact/threaten what they think is important.
- Privacy: Risks from the social pitfalls associated with the use and misuse of an individual's data.
- Social Justice and Equity: Risks from business practices and technologies that marginalize or disadvantage specific segments within society.
- Social Trends: Risks from shifts in social norms, changing consumer expectations, or evolving cultural behaviors.
- Worldview: Risks from people's deeply held beliefs about how they view the world and how it should function.

Applying a Risk Innovation Approach to Advanced Brain Machine Interfaces

The risk innovation framework we are developing is particularly pertinent to the advanced brain machine interface technology outlined by Musk et al [4]. Although the technological advances presented do not explicitly address the pathway between these developments and their successful application, a risk innovation approach provides insights into where there are potential threats to value—or development pinch-points—that may arise, and early actions that may be taken to navigate these. Importantly, it opens the way to win-win scenarios between technological functionality, commercial success, and ethical and socially responsible innovation.

Based on the technology described, we explore below how risk innovation may be used to provide a starting point for thinking through the potential challenges and opportunities surrounding the ethical and socially responsible development of the technology. This application is necessarily limited by assumptions around the development and use of the technology. However, it nevertheless provides novel insights into the technology's development and enables questions associated with ethical, responsible, and successful innovation to be bound in ways that potentially reveal useful pathways forward.

As Musk et al [4] articulated, there is clear value in the advanced brain machine interface systems they describe for therapeutic use. This is a technology that has the potential to provide novel approaches for addressing a spectrum of neurological conditions, from Parkinson disease and epilepsy to debilitating migraines. It could also vastly improve how users are able to interface with and control prosthetics, including those that replace diminished sensory function.

There is also the expectation of the technology being used to enhance performance in healthy users. Although this is not articulated in the paper, Elon Musk and Neuralink have been explicit about the possibilities of using technology to augment the capabilities of users through high-speed, wireless brain machine interfaces [21]. Here, apps could be as diverse as aimed at gaming and cognitive enhancement or sensory enrichment.

These potential capabilities translate into the hypothesized (and not exclusive) areas of value for the enterprise, investors, customers, and communities that are listed in Textbox 2. These are indicative only, but illustrate how areas of value potentially vary between different groups. For instance, investors are likely to be interested in how others perceive the trustworthiness of developers as a proxy for long-term return on investment. Customers are likely to value the degree to which brain machine interfaces are secure from interference and ensure that they, as users, retain autonomy over their actions and brain function. On the other hand, communities potentially touched by the technology are likely to place a high premium on people with brain machine interfaces becoming somehow "other" and a threat to social norms and expectations.

Comparing the areas of value in Textbox 2 and the orphan risks listed in Textbox 1, it is possible to map relevant risks onto key areas of value. This exercise is necessarily subjective. However, it begins to provide novel insights into potential challenges worth exploring further if key areas of value are to be protected and nurtured.

Figure 1 provides an example of what this metaphorical "risk landscape" might look like for the brain machine interface, as described by Musk et al [4]. The approach is designed to provide a snapshot of areas that warrant further attention and to illuminate potential risk clusters that may otherwise remain obscured. It is not inclusive or comprehensive and does not include many conventional risks for which there are established risk assessment and mitigation frameworks, for example, cyber security. Rather, it provides a starting point for mapping out orphan risks that have the potential to directly or indirectly affect development and that could raise substantial issues around ethical and socially responsible innovation.

Here, it should be emphasized that this mapping exercise is qualitative and designed to focus and constrain perspectives on potentially blindsiding risks while encouraging innovative thinking around potential barriers to success. The methodology is intended to be iterative and open up new possibilities, without being prescriptive.

Within this context, the mapping presented in Figure 1 can be interpreted in a number of ways. First, it provides insights into orphan risks that are worth the enterprise being aware of. In the case of advanced brain machine interfaces, how investors, users, and others perceive the potential workings and impacts of the technology are flagged as important, as are social trends that may either create opportunities for development or potential barriers—for instance, if there is a public backlash against brain machine interface-based augmentation. Perhaps, not surprisingly, for an invasive neural read-write technology, novel health impacts are flagged, as are fundamentally unpredictable "black swan" events. In this case, the mapping suggests a high level of attention should be paid to enterprise and technology agility with respect to navigating around unexpected hurdles.

Textbox 2. Hypothesized areas of value associated with advanced brain machine interfaces. These are used as an example of how a risk innovation approach can help map out and navigate a complex risk landscape. They do not necessarily reflect areas of value as defined by the developers of advanced brain machine interfaces. The textbox is intentionally limited to three areas of value per column in order to avoid paralysis by overanalysis.

Enterprise

- Transformational medical interventions
- Low cost, highly accessible brain machine interface–enabled performance enhancement
- Technological leadership

Investors

- Products that deliver on their promise
- Brand trustworthiness
- High return on investment

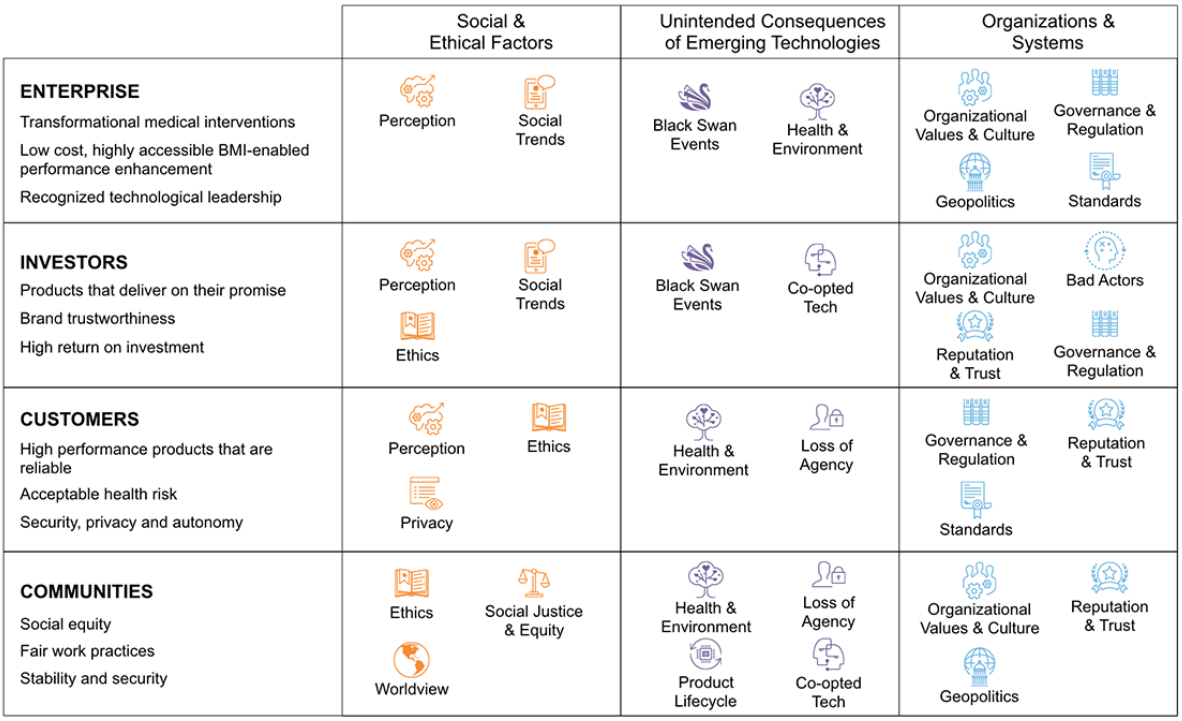
Consumers

- High performance products that are reliable
- Acceptable health risk
- Security, privacy, and autonomy

Communities

- Social equity
- Fair work practices
- Stability and security

Figure 1. A qualitative snapshot of the possible orphan risk landscape associated with advanced brain machine interfaces.



Reading across the map, there is a greater density of orphan risks identified under "Organizations and Systems," suggesting that some of the greatest potential threats to successful advanced brain machine interface technologies lie within the enterprise and the formal organizations governing how it operates. These include the degree to which organizational values and culture align with how the technology is developed and used, evolving regulations that may play an outsized influence on development pathways, and national and international standards that guide and limit technology performance and use.

Second, the risk landscape in Figure 1 provides insights into potential threats to value within key constituencies, which may, in turn, lead to barriers to success. Here, orphan risks that do not directly impact the enterprise, but are likely to influence its success, are highlighted. These include the ethics surrounding how the technology is developed and used, how it potentially leads to social injustice, and whether it raises privacy concerns. This broader perspective on orphan risks also flags possible issues such as how trustworthy the enterprise is perceived to be and the dangers of "bad actors" giving the technology a negative reputation that risks poisoning the market.

A third way the landscape in Figure 1 can be read is by identifying risks that cluster and converge in ways that increase the chances of truly blindsiding impacts. Certain risks dominate the landscape, including those associated with ethics, health, governance, and regulation. If not planned for, these could build upon each other to create insurmountable obstacles to value creation. Instead of planning for just one risk at a time, analysis and understanding of these clusters have the potential to provide insights into where resources are best focused to protect stakeholder value. Here, the analysis suggests that some of the more substantial near-time challenges to ensuring successful, responsible, and ethical advanced brain machine interfaces are associated with organizational systems and operations. These include internal processes such as the development and application of guiding principles and expectations of ethical and responsible behavior. However, they also extend to professional bodies, regulatory agencies, and communities of practice that form the broader ecosystem within which advanced brain machine interfaces are conceived, developed, and used. Here, Figure 1 also highlights orphan risks that occur multiple times such as ethics, perception, health and environmental impacts, and reputation and trust. These (and similar risk clusters) provide insights into where particular attention is most likely needed to support successful and responsible development.

Summary

Applying a risk innovation methodology to the brain machine interface technology described by Musk et al [4] is admittedly subjective, especially as their paper primarily focuses on technological advances, and not on how they may be fully implemented into products and services in the future. However, its application provides a unique framework for exploring the ethical and socially responsible development of the technology in ways that are bounded by pragmatic considerations and are neither misguided by myopic optimism nor stymied by overspeculative pessimism. The approach taken here is intentionally focused on identifying pathways to innovation success and can thus be seen to favor the enterprise. Yet, by broadening the understanding of risk to key constituencies that are potentially impacted by the enterprise and the technology, it provides a canvas on which successful innovation may be aligned with ethical and socially responsible innovation. We would go so far as to say that it provides a conceptual framework and methodology that enables innovators to create and grow value by being intentionally responsive to ethical and social issues.

In the case of the brain machine interface technology described by Musk et al [4], this approach indicates that there is a crowded risk landscape between the current state of the technology and its successful development and use, with many of these risks pivoting on how companies address potentially serious ethical and social issues. These include the potential for the technology to widen the gap between the rich and poor (and the privileged and the marginalized); to challenge social norms around the use of enhancement technologies; and to raise complex issues around data privacy, user security, and autonomy—especially where manufacturers retain ownership of implanted brain machine interfaces or their operation and upkeep is dependent on a subscription service for instance. They also highlight the need for developers and others to consider with some seriousness how they build trust with the communities they depend on and actions that may erode trust.

By mapping out the orphan risk landscape they face—and iterating frequently as the landscape shifts and changes—there is no reason why the technical developments beginning to emerge around advanced brain machine interfaces will not lead to powerful therapeutic interventions and even transformative enhancement capabilities, which are economically successful, ethical, and socially responsible. Yet, this will only come about if developers are capable of looking beyond technical performance and conventional risks, culturing a sophisticated understanding of the risk landscape they face and the approaches they need to adopt in order to successfully navigate it.

The risk innovation approach we have described provides novel insights into this landscape and is a useful tool for revealing potentially blindsiding risks while there is still time to take corrective action. In the case of advanced brain machine interfaces, it helps map out challenges and opportunities that may otherwise be easy to miss, but have the ability to derail progress. Yet, it is just the start of a journey toward developing products that are capable of changing lives for the better, without causing substantial harm. Here, success depends on the willingness of innovators and others to take ethical and responsible innovation seriously and to draw on the interdisciplinary and transdisciplinary expertise that is necessary to translate good intentions into positive outcomes. Here, our hope is that this is a pathway that advanced brain machine interface technologies will follow.

Conflicts of Interest

None declared.

References

1. Wilson BS, Dorman MF. Cochlear implants: A remarkable past and a brilliant future. Hearing Research 2008 Aug;242(1-2):3-21. [doi: 10.1016/j.heares.2008.06.005]
2. Perlmutter JS, Mink JW. Deep brain stimulation. Annu Rev Neurosci 2006 Jul 21;29(1):229-257 [FREE Full text] [doi: 10.1146/annurev.neuro.29.051605.112824] [Medline: 16776585]
3. Maynard EM, Nordhausen CT, Normann RA. The Utah Intracortical Electrode Array: A recording structure for potential brain-computer interfaces. Electroencephalography and Clinical Neurophysiology 1997 Mar;102(3):228-239. [doi: 10.1016/s0013-4694(96)95176-0]
4. Musk E, Neuralink. An Integrated Brain-Machine Interface Platform With Thousands of Channels. J Med Internet Res 2019 Oct;21(10) [FREE Full text] [doi: 10.2196/16194] [Medline: 31642810]
5. Maguire GQ, McGee EM. Implantable Brain Chips? Time for Debate. The Hastings Center Report 1999 Jan;29(1):7. [doi: 10.2307/3528533]
6. Lucivero F, Tamburrini G. Ethical monitoring of brain-machine interfaces. AI & Soc 2007 Aug 3;22(3):449-460. [doi: 10.1007/s00146-007-0146-x]
7. Burwell S, Sample M, Racine E. Ethical aspects of brain computer interfaces: a scoping review. BMC Med Ethics 2017 Nov 09;18(1):60 [FREE Full text] [doi: 10.1186/s12910-017-0220-y] [Medline: 29121942]
8. Rommelfanger K, Jeong S, Ema A, Fukushi T, Kasai K, Ramos K, et al. Neuroethics Questions to Guide Ethical Research in the International Brain Initiatives. Neuron 2018 Oct;100(1):19-36. [doi: 10.1016/j.neuron.2018.09.021]
9. Keskinbora KH, Keskinbora K. Ethical considerations on novel neuronal interfaces. Neurol Sci 2017 Dec 2;39(4):607-613. [doi: 10.1007/s10072-017-3209-x]
10. The RS. iHuman: blurring lines between mind and machine. London: The Royal Society; 2019. URL: https://royalsociety.org/-/media/policy/projects/ihuman/report-neural-interfaces.pdf [accessed 2019-09-16]
11. Maynard AD. Why we need risk innovation. Nature Nanotech 2015 Sep 3;10(9):730-731. [doi: 10.1038/nnano.2015.196]
12. Risk Innovation Nexus. URL: http://riskinnovation.org [accessed 2019-09-15]
13. Maynard AD, Garbee E. Responsible innovation in a culture of entrepreneurship: A US perspective. In: von Schomberg R, Hankins J, editors. International Handbook on Responsible Innovation: A Global Resource. Cheltenham, UK: Edward Elgar; 2019.
14. Mohorčich J, Reese J. Cell-cultured meat: Lessons from GMO adoption and resistance. Appetite 2019 Dec;143:104408. [doi: 10.1016/j.appet.2019.104408]
15. Maynard AD. The (nano) entrepreneur's dilemma. Nature Nanotech 2015 Feb 5;10(3):199-200. [doi: 10.1038/nnano.2015.35]
16. Guston DH. Understanding 'anticipatory governance'. Soc Stud Sci 2013 Nov 15;44(2):218-242. [doi: 10.1177/0306312713508669]

17. World Economic Forum. Agile Governance. Reimagining Policy-making in the Fourth Industrial Revolution. Geneva: World Economic Forum; 2018. URL: http://www3.weforum.org/docs/WEF_Agile_Governance_Reimagining_Policy-making_4IR_report.pdf [accessed 2019-09-16]
18. von Schomberg R, Hankins J, editors. International Handbook On Responsible Innovation: A Global Resource. Cheltenham, UM: Edward Elgar Pub; 2019.
19. Shane S, Metz C, Wakabayashi D. How a Pentagon Contract Became an Identity Crisis for Google.: New York Times; 2018 May 30. URL: https://www.nytimes.com/2018/05/30/technology/google-project-maven-pentagon.html [accessed 2019-09-16]
20. Maynard AD. Edge of Innovation.: Medium; 2018 Dec 12. It's time for tech startups and their funders to take “orphan risks” seriously URL: https://medium.com/edge-of-innovation/its-time-for-tech-startups-and-their-funders-to-take-orphan-risks-seriously-3f7813976a07 [accessed 2019-09-16]
21. Neuralink. URL: https://www.neuralink.com/ [accessed 2019-09-16]

Edited by G Eysenbach; submitted 18.09.19; peer-reviewed by J Salisbury; comments to author 12.10.19; revised version received 13.10.19; accepted 14.10.19; published 31.10.19.

Please cite as:
Maynard AD, Scragg M
The Ethical and Responsible Development and Application of Advanced Brain Machine Interfaces
J Med Internet Res 2019;21(10):e16321
URL: http://www.jmir.org/2019/10/e16321/
doi:10.2196/16321
PMID:31674917

Commentary

The Reconnecting the Hand and Arm with Brain (ReHAB) Commentary on "An Integrated Brain-Machine Interface Platform With Thousands of Channels"

Robert F Kirsch[1*], BSc, MSc, PhD; A Bolu Ajiboye[2*], BSc, MSc, PhD; Jonathan P Miller[2*], BSc, MD

[1]VA Rehabilitation Research and Development Service, Center for Functional Electrical Stimulation, Case Western Reserve University, Cleveland, OH, United States

[2]Case Western Reserve University, University Hospitals Cleveland Medical Center, Cleveland, OH, United States

[*]all authors contributed equally

Corresponding Author:
Robert F Kirsch, BSc, MSc, PhD
VA Rehabilitation Research and Development Service, Center for Functional Electrical Stimulation
Case Western Reserve University
10900 Euclid Avenue
Wickenden Building 340
Cleveland, OH, 44106-7207
United States
Phone: 1 2163683158
Email: rfk3@case.edu

Related Article:

Comment on: http://jmir.org/2019/10/e16194/

Abstract

Intracortical brain-machine interfaces are a promising technology for allowing people with chronic and severe neurological disorders that resulted in loss of function to potentially regain those functions through neuroprosthetic devices. The penetrating microelectrode arrays used in almost all previous studies of intracortical brain-machine interfaces in people had a limited recording life (potentially due to issues with long-term biocompatibility), as well as a limited number of recording electrodes with limited distribution in the brain. Significant advances are required in this array interface to deal with the issues of long-term biocompatibility and lack of distributed recordings. The Musk and Neuralink manuscript proposes a novel and potentially disruptive approach to advancing the brain-electrode interface technology, with the potential of addressing many of these hurdles. Our commentary addresses the potential advantages of the proposed approach, as well as the remaining challenges to be addressed.

(J Med Internet Res 2019;21(10):e16339) doi:10.2196/16339

KEYWORDS

brain computer interfacing; intracortical recording; neural engineering; neurosurgery

Over the past two decades, several academic research laboratories have advanced the science and implementation of intracortical brain-machine interfaces from exclusive use in nonhuman primates to investigative use in human volunteers with chronic neurological impairments under pilot clinical trials approved through the US Food and Drug Administration (FDA). Researchers in these labs have successfully demonstrated that these persons can use intracortical brain-machine interfaces to command movements of a computer cursor, a robotic arm, their own arm reanimated through functional electrical stimulation, and recently speech articulation. Our own group has developed a system that uses an intracortical brain-machine interface to command a functional electrical stimulation system that coordinates activation of paralyzed upper limb muscles to restore useful function. Our current study is the Reconnecting the Hand and Arm with Brain (ReHAB) clinical trial, which uses Blackrock recording and stimulation arrays.

As noted in the paper by Musk and Neuralink [1], the sampling of neurons by electrode arrays currently available for human applications is a tiny percentage of the relevant neural population. Access has been largely limited to cortical surfaces

1-2 millimeters deep and on gyri and not in sulci, which is where several key areas in the human brain are located. The performance of existing human grade electrodes deteriorates over a shorter time frame than would be acceptable for many clinical applications. Transmitting the recorded neural signals to the external world with high fidelity has relied on percutaneous connectors that are unlikely to provide permanent clinical solutions, or on low channel–count telemetry devices that are unlikely to provide adequate information for many applications. In addition to recording, intracortical stimulation through these same or similar electrodes is being tested to inject information into brain structures (eg, to restore sensation in paralyzed individuals). Intracortical stimulation applications face many of the same issues described above for recording but bring additional safety issues. The Musk and Neuralink paper hints at future brain stimulation capabilities, but these are not described and will not be discussed further in our commentary.

Based on our experiences, any significant step forward in human brain recording will require electrode technologies that provide orders of magnitude more information. This information will be derived from: (1) a much higher number of recording/stimulation contacts; (2) recording both surface and deep brain structures (eg, in sulci or noncortical regions), as well as from multiple brain areas with different functionality; and (3) different signal types (eg, single units and local field potentials). These electrodes should be able to be inserted safely and within a reasonable surgical time window, and should be well tolerated by the brain (ie, cause negligible damage upon insertion and be virtually invisible to the immune system to thus provide stable performance over decades). Hardware that records signals from these electrodes should extract relevant neural information from the high numbers of channels with high fidelity and in real-time, avoid percutaneous interfaces, and always be available for use.

The authors of this paper described a novel approach involving ultrafine polymer threads, each containing a dense linear array of 32 high-impedance electrodes, that are individually implanted into the cortex using a robotic device. The implantation robot can be controlled intraoperatively with micron precision to avoid small surface blood vessels and is able to implant up to six threads (192 electrodes) per minute. A small (23x18.5 millimeter) custom printed circuit board with onboard power is able to connect to up to 96 of the threads for a total of 3072 electrodes per implanted array, and digitized high-bandwidth neural data is streamed from the device using a single USB-C cable. Two different versions of the system were developed, one maximizing reliable manufacturing (using half as many leads) and another maximizing channel count. Testing in an awake rat demonstrated the ability to record signals interpreted as neural spike data from 43% of the channels. The authors conclude that this strategy may revolutionize brain-machine interfaces by recording from an unprecedented number of neurons.

Clinical application of brain-machine interfaces may require a very high channel count to allow recording from many neurons, so the relatively small number of contacts available using currently available platforms has hindered its development. The technology and processes presented in the Musk and Neuralink paper can certainly increase the channel count by an order of magnitude, providing a more detailed sampling of relevant signals as well as some welcome redundancy. Perhaps even more attractive about the proposed approach is the possibility of placing electrodes into areas of the brain that have been difficult or impossible to reach with existing intracortical arrays (eg, on the medical surfaces of sulci and subcortical structures). These new locations could potentially provide different types of information for enhancing brain-machine interface performance (eg, abstract planning of activities [including movement goals and sequencing], anticipated reward signals, and decision making), as well as better elucidate the interactions between different brain structures (eg, between different motor areas, processing of sensory information, and integration of sensory-motor activities). Such information is likely to be critical for successful performance in more complex brain-machine interface applications, such as multidimensional arm movement involving complex physical dynamics.

The use of individually implanted threads of electrodes is a clever way to allow an exponential increase of channel count over existing brain-machine interface technology, and the authors' approach is well conceived with a good consideration of both physical and technical characteristics. However, the potential clinical application of this strategy is unclear since it has only been tested in a small number of rodents, with no comparison to existing approaches or verification of safety using histological analysis after implantation. The authors claim that their implants will have greater longevity than other options because of less immune response related to electrode stiffness and microvascular disruption, but no evidence is presented to support either of these assumptions, and improved durability was not verified using long-term implantation. It is not clear that blood vessels below the surface can be avoided, potentially critical for immune responses. The paper does not address the use of the thread electrodes for larger brains with more complex cortical structures (eg, the deeply folded structure of the human brain). The potentially implantable recording system as presented does not include hermetic sealing, a relevant power source (eg, battery, induction, or optical), or a technique (eg, wireless) for transmitting high bandwidth data out of the body without a percutaneous interface. Furthermore, the potential impact of noise and artifact has not been unequivocally established: It is uncertain whether the signals recorded in the rodent study actually represent meaningful neural data since the measured impedance was relatively low compared with penetrating electrodes typically used to measure single-unit neural activity, similar signals were seen on many adjacent channels, and no attempt was made to validate that the spike data had physiological characteristics typical of neural spiking patterns in the regions that were implanted. The technology is very innovative, but better validation will be necessary to establish its clinical potential.

Overall, this new technology is exciting because it attempts to directly address a number of bottlenecks that have hindered true clinical translation of intracortical arrays for use in brain-machine interfaces. The amount of high-resolution information that could be simultaneously recorded from cortical neurons may lead to new advances in the application of data

mining and machine learning approaches to better understand cortical electrophysiological activity at the macro-, meso-, and microscale levels. Ultimately these multi-prong advances will likely result in advances in brain-controlled neuroprosthetics for addressing impaired function in neurologically compromised individuals. However, a responsible stance of cautious optimism must be taken. There is a long road between showing single neuron recordings in a handful of rodents to clinically translated human use, including proof (not just potential) of longevity, efficacy, and safety through an FDA-approved human clinical trial. If successful, the proposed technology (and future derived technologies) could pave the way for more widespread translation of intracortical brain-machine interfaces for medical applications in people with chronic neurological impairments.

Conflicts of Interest

None declared.

Reference

1. Musk E, Neuralink. An Integrated Brain-Machine Interface Platform With Thousands of Channels. J Med Internet Res 2019 Oct;21(10) [FREE Full text] [doi: 10.2196/16194] [Medline: 31642810]

Abbreviations

FDA: US Food and Drug Administration
ReHAB: Reconnecting the Hand and Arm with Brain

Edited by G Eysenbach; submitted 19.09.19; peer-reviewed by J Pancrazio, M Wang; comments to author 08.10.19; revised version received 15.10.19; accepted 17.10.19; published 31.10.19.

Please cite as:
Kirsch RF, Ajiboye AB, Miller JP
The Reconnecting the Hand and Arm with Brain (ReHAB) Commentary on "An Integrated Brain-Machine Interface Platform With Thousands of Channels"
J Med Internet Res 2019;21(10):e16339
URL: http://www.jmir.org/2019/10/e16339/
doi:10.2196/16339
PMID:31674921

Commentary

From Novel Technology to Novel Applications: Comment on “An Integrated Brain-Machine Interface Platform With Thousands of Channels” by Elon Musk and Neuralink

Alexander N Pisarchik[1,2*], PhD; Vladimir A Maksimenko[2*], PhD; Alexander E Hramov[2*], PhD, SD

[1]Center for Biomedical Technology, Technical University of Madrid, Pozuelo de Alarcón, Madrid, Spain

[2]Neuroscience and Cognitive Technology Laboratory, Center for Technologies in Robotics and Mechatronics Components, Innopolis University, Innopolis, Russian Federation

[*]all authors contributed equally

Corresponding Author:
Alexander N Pisarchik, PhD
Center for Biomedical Technology
Technical University of Madrid
Campus Montegancedo
Pozuelo de Alarcón, Madrid, 28223
Spain
Phone: 34 910679279
Email: alexander.pisarchik@ctb.upm.es

Related Article:

Comment on: http://jmir.org/2019/10/e16194/

(J Med Internet Res 2019;21(10):e16356) doi:10.2196/16356

KEYWORDS

brain-computer interface; brain-machine interface; brain; electroencephalography

The first attempts to translate neuronal activity into commands to control external devices were made in monkeys yet in 1960s [1]. After that, during 1960-1970, the biological feedback was realized in monkeys, to provide voluntary control of the firing rate of cortical neurons [2,3]. The term “brain-computer interface” appeared only in earlier 1970s [4]. The brain-computer interface is usually referred to as a “brain-machine interface” in invasive studies. Nowadays, the brain-computer interface and brain-machine interface research and applications are considered one of the most exciting interdisciplinary areas of science and technology.

In particular, brain-computer interfaces are very promising for neurorehabilitation of sensory and motor disabilities [5], neurocommunication [6], exoskeletons [7], cognitive state evaluation [8], etc. Advanced mathematical methods for extraction and classification of neuronal activity features hold out hope for the future use of brain-computer interfaces in everyday life. At the same time, the lack of effective invasive neuroimaging techniques providing a high-resolution neural activity recording for medical purposes limits the brain-machine interface implementation in clinics.

In their paper, Elon Musk and Neuralink [9] have successfully addressed the major issues hampering the next generation of invasive brain-computer interface (or brain-machine interface) development by introducing a novel integrated platform enabling a high-quality registration of thousands of channels. Their device contains arrays of flexible electrode threads with up to 3072 electrodes per array, distributed across 96 threads. To overcome a surgical limitation, the authors have built a neurosurgical robot that inserts 6 threads per minute with a micrometer spatial precision. To increase the biocompatibility, they created a neurosurgical robot, which implants polymer probes much faster and more safely than existing surgical approaches. Using this platform in freely moving rats, the authors report a spiking yield of up to 85.5%.

Although the developed system is considered an effective platform for research in rodents, it can serve as an invasive neurointerface prototype for clinical applications. Specifically, multielectrode neurointerfaces may become the basis for new communication systems and advanced assistive technologies for paralyzed people as well as control external devices and interact with the entire environment, eg, by integrating into new fast developed technologies, such as Smart Home and Internet of Things. Moreover, the brain-computer interface applications

are very promising for detecting hidden information in the user's brain, which cannot be revealed by conventional communication channels. Currently, the use of noninvasive brain-computer interfaces in these fields is limited by a low number of commands that can be recognized. This limitation arises from a relatively small number of features, which can be extracted from the scalp-level electroencephalography or functional near-infrared spectroscopy recordings. The invasive brain-computer interfaces (or brain-machine interfaces) demonstrate much better performance than noninvasive brain-computer interfaces; however, they require a larger number of channels to obtain more detailed information about the individual spiking activity of the neurons across distributed cortical regions. The device reported in the paper of Elon Musk and Neuralink approaches the solution to this problem.

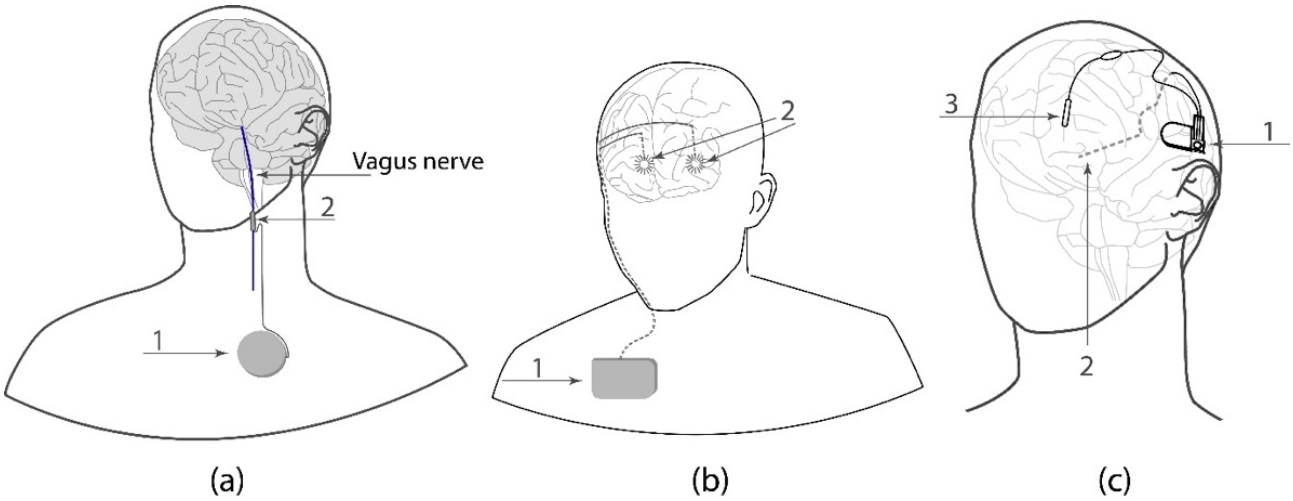
One of the most promising applications of noninvasive brain-computer interfaces is monitoring, control, and training of human's psychophysiological states and cognitive abilities. In such studies, the subject's mental state is continuously evaluated by the passive brain-computer interface. The passive brain-computer interface analyzes the current brain activity of the user without any aim to control command generation and provides information about features of the actual brain activity related to attention, emotional state, fatigue, etc [10]. These top-down processes originating on the cortical level are spatially distributed and relatively slow, and therefore, they have well-pronounced markers on the noninvasive electroencephalography and functional near-infrared spectroscopy signals.

However, in the case of spatially localized neuronal activity, the noninvasive techniques are unable to recognize distinguished features of neurophysiological diseases in real time, and therefore, one needs to insert electrodes into particular brain regions. For example, to predict epileptic seizures, the neuronal activity must be recorded from predefined focal areas of the brain, where an earlier manifestation of this pathological activity is mostly pronounced and can be promptly detected in real time [11]. Recently, an efficient method for epilepsy prediction based on electrical brain activity was proposed [12,13]. The method allows forecasting a focal seizure up to 5 seconds before it occurs. This time is sufficient for a brain-machine interface to generate a proper signal, which suppresses the forthcoming seizure. For drug-resistant patients, the complete abolishment of epileptic seizures might be achieved by a brain-machine interface/brain-computer interface that predicts a seizure onset, combined with a system that interferes with the process that causes the seizure. Thus, seizure prediction remains an unresolved problem due to insufficient information about neural processes in the onset brain area. Therefore, one possible and very important clinical application of the Neuralink technology is a brain-machine interface for patients with drug-resistant epilepsy. These brain-machine interfaces should imply the brain stimulation (electrical, magnetic, optogenetic, etc) to interrupt or even prevent epileptic seizures. The stimulation can be delivered to the brain in either an open-loop or a closed-loop fashion. In the former case, there is no need to monitor the current brain activity, since the stimulation is activated manually or in accordance with a predefined stimulation protocol.

As an example of the open-loop system, in Figure 1a,we present the vagus nerve stimulator manufactured by Cyberonics, Inc (Texas, United States) in 1977. The stimulator contains an implantable pulse generator and an electrode to stimulate the left vagus nerve in a repetitive "duty cycle" ("on" for 30 seconds and then "off" for 5 minutes), which allows reducing the number of seizures by an average of 30%–40% [14]. Along with the vagus nerve stimulator, other open-loop antiepileptic devices are also used for deep brain stimulation. One of the first open-loop deep brain stimulators was the stimulator of the anterior nucleus of the thalamus in epilepsy proposed by Medtronic, Inc (United States) for patients with partial-onset epilepsy [15] (Figure 1b).

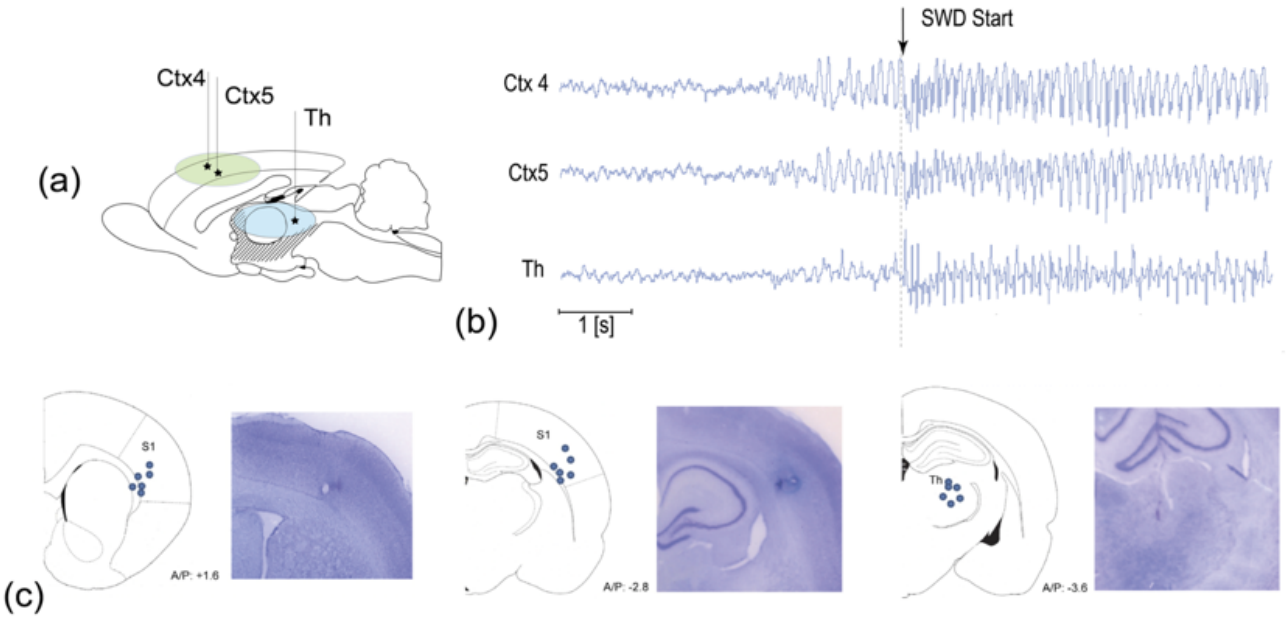
Although the open-loop systems enable a significant reduction in the number of epileptic seizures, it is obvious that a more efficient control of epileptic activity requires continuous monitoring of the current brain activity, which can be achieved by using closed-loop brain-machine interfaces. One of the first closed-loop brain-machine interfaces was the responsive neurostimulator designed by Neuropace Inc. (California, United States; Figure 1c). It contains implanted electrodes for recording the intracranial electroencephalography used by the algorithm, which determines the moment of time when a seizure starts. To interrupt the seizure, the triggered focal electrical stimulation is sent to a specific brain area [16].

Figure 1. The schematic illustration of brain-machine interface prototypes to suppress epileptic seizures using electrical stimulation. (a) Vagus nerve stimulator containing (1) an implantable pulse generator and (2) a stimulation lead. (b) Stimulator of the anterior nucleus of the thalamus in epilepsy containing (1) an implantable pulse generator and (2) intracranial electrodes placed in the anterior thalamic nuclei bilaterally. (c) Responsive neurostimulator containing (1) implanted deep electrodes for recording electroencephalography signals, (2) an implantable device for processing electroencephalography signals from electrodes, and (3) strip electrodes receiving an electrical stimulation signal generated by the device to stop seizures.



It is important to point out that epileptic seizures are well detected using electrocorticography or intracranial electroencephalography, which display a pronounced marker of the high-amplitude rhythmic activity. Recent studies reported a 100% accuracy of this technique in detecting epileptic seizures in rats [17]. Since the preictal activity may not differ from a normal behavior, the prediction of seizures is a very challenging task. Although existing algorithms allow seizure predictions with high sensitivity, they are too complicated and too specific to be used in clinics [18]. One of the closed-loop systems for the seizure prediction and prevention was recently tested in vivo in rats [11] (Figure 2). The seizure prediction algorithm was based on the electrocorticography signals recorded by three electrodes in the cortex and the thalamus, as shown in Figure 2a. The brain-machine interface was able to correctly predict 45% seizures, but the number of false predictions varied from 20 to 100 per hour among animals. In this regard, one can expect that the technique developed by Elon Musk and Neuralink with thousands of channels will significantly improve seizure prediction. The large number of registered channels will increase classification accuracy in pre-epileptic state recognition and decrease the number of false positives during light slow wave sleep. This is indeed a step toward the next generation of brain-machine interfaces for drug-resistant epilepsy. The accurate seizure prediction algorithm based on the multichannel Neuralink technology enables the prevention of ongoing seizures and protects the patient against unnecessary stimulations caused by false alarms. Thanks to the developed brain-machine interface platform with thousands of channels, significant progress is expected in solving this important problem, enabling new clinical trials for patients resistant to drug therapy.

Figure 2. (a) Schematic representation of the experiments with a rat. (b) The set of electrocorticography recordings taken from subgranular layers 4 (Ctx4) and 5 (Ctx5) of the somatosensory cortex and postero/lateral thalamus before and during onset of the epileptic spike-wave discharge. (c) Histological verification of the electrode location in the somatosensory cortex (S1) and postero/lateral thalamus (based on data from [<xref ref-type="bibr" rid="56ref11">11</xref>]). Th: thalamus.



Another type of brain dynamics associated with spatially localized cortical sources is a motor-related neuronal activity. This problem has a significant social impact. Today, most brain-computer interfaces and brain-machine interfaces are designed for patients with severe motor disorders. The brain-computer interface allows a disabled person to control wheelchairs, exoskeletons, and robotic manipulators by generating commands via voluntary changes of brain activity in the motor cortex induced by motor imagery. For instance, these brain-computer interfaces are used to control a cursor or a wheelchair in two dimensions with the intention of moving the left or right hand [19]. However, the performance of such noninvasive brain-computer interfaces for movement control is limited by a small number of commands. Studies on animals, primarily monkeys, have shown the effectiveness of invasive registration of cortical neuron activity to create an effective brain-machine interface for controlling more complex movements. In 2008, an invasive interface was implemented, which allows a monkey to control an anthropomorphic manipulator [20]. In the experiment, the signal from 15-25 cortical units was recorded in the monkey, to control a robotic arm to feed itself. The monkey performed a continuous self-feeding task with a mean success rate of 69.5%. Similar results were obtained with an invasive interface that controlled the lower limbs [21] or both hands simultaneously (bimanual movements) [22] using cortical activity patterns.

It should be noted that the use of invasive brain-machine interfaces allows patients with tetraplegia to perform reach and grasp tasks with a robotic arm manipulator [23]. At the same time, the brain-machine interface implementation in humans is still not employed in clinical practice due to surgical difficulties and the problems of biocompatibility. In this context, the proposed neurosurgical robot can be considered an important step toward human implants. A significant advantage of the robot is its rapid manipulation to insert six-electrode threads per minute in addition to its capacity in precisely inserting the most biocompatible polymer probes. Another potential clinical application of the brain-machine interface might be the restoration of neuronal connections, lost due to degenerative diseases, such as Alzheimer disease, or replacement of dead neurons with artificial ones. Comprehensive studies covering many aspects of this issue are currently underway [24].

Finally, we would like to draw attention to the importance of the neural activity modulation in the next-generation brain-machine interfaces. This possibility is essential for applications such as neuroprosthetics in order to provide biological feedback as a sense of touch [25], and new therapeutic approaches for patients with drug-resistant epilepsy in order to prevent ongoing seizures by the appropriate electric stimulation [26]. In view of the foregoing, the authors of the paper highlight the capability of their custom electronics to deliver an electrical stimulation to every channel, but unfortunately, do not present these results in the paper. It would be interesting to know how the authors plan to deliver electrical pulses to cells and record the neural activity simultaneously, in particular, whether the stimulation preserves the potential for simultaneous recording of the neural activity with the minimized effect of stimulus-induced artifacts. If this issue is successfully addressed, it would become possible to interact with the neural activity in the continuous manner instead of the neuron stimulation with the electrical pulses in a discrete closed-loop fashion. Nowadays, this is achieved by the optogenetic brain stimulation [27,28] with the help of hybrid optoelectronic interfaces [29] and thought to be a substantial advantage of optogenetics over the electrical stimulation and recording.

One of the main obstacles for clinical applications of implanted devices is their low biocompatibility that does not allow long-term recordings. Elon Musk and Neuralink approach this problem by utilizing a biocompatible polyimide, which encapsulates a gold thin film trace. By choosing electrically

conducted materials, one has to play between impedance and biocompatibility. The authors tested the polyethylenedioxythiophene-doped polymer with polystyrene sulfonate and iridium oxide and achieved lower impedance for the former, but better biocompatibility for the latter. They promise to continue research in this direction and extend their techniques and processes to other types of conductive electrode materials and coatings.

Last, but not the least, are the unpleasant consequences of this work that we should pay attention to, since every achievement of humanity has two sides. On one hand, it intends to improve the quality of life, but on the other hand, it can be used by unscrupulous people for their selfish goals. Therefore, every scientist should think not only about a positive impact of his/her research, but also about its possible negative effects. Among the undesired effects of brain-machine interfaces with electrodes implanted into the human brain is the possibility that a government or a nongovernmental organization will control and manipulate the person's behavior not only through mass media, but also by directly sending commands to the brain. In this regard, numerous debates about ethics of using brain-machine interfaces are currently underway in the media.

In summary, the novel neurointerface by Elon Musk and Neurolink has all chances to become a real step forward to the next generation of brain-machine interfaces for both research and clinical applications. Invasive interfaces can help disabled people control external devices and communicate with other people. We believe that future communication technologies will be based on brain-computer interfaces that will read brain signals and translate them to messages, which then will be sent to mobile or other devices. Furthermore, invasive brain-machine interfaces will allow a direct communication between people by their thoughts.

Acknowledgments

This work has been supported by the Center for Technology in Robotics and Mechatronics Components (Innopolis University).

Conflicts of Interest

None declared.

References

1. Evarts EV. Pyramidal tract activity associated with a conditioned hand movement in the monkey. J Neurophysiol 1966 Nov;29(6):1011-1027. [doi: 10.1152/jn.1966.29.6.1011] [Medline: 4961643]
2. Fetz EE. Operant conditioning of cortical unit activity. Science 1969 Feb 28;163(3870):955-958. [doi: 10.1126/science.163.3870.955] [Medline: 4974291]
3. Fetz EE, Finocchio DV. Operant Conditioning of Specific Patterns of Neural and Muscular Activity. Science 1971 Oct 22;174(4007):431-435. [doi: 10.1126/science.174.4007.431]
4. Vidal JJ. Toward direct brain-computer communication. Annu Rev Biophys Bioeng 1973;2:157-180. [doi: 10.1146/annurev.bb.02.060173.001105] [Medline: 4583653]
5. Daly JJ, Wolpaw JR. Brain-computer interfaces in neurological rehabilitation. Lancet Neurol 2008 Nov;7(11):1032-1043. [doi: 10.1016/S1474-4422(08)70223-0] [Medline: 18835541]
6. Maksimenko VA, Hramov AE, Frolov NS, Lüttjohann A, Nedaivozov VO, Grubov VV, et al. Increasing Human Performance by Sharing Cognitive Load Using Brain-to-Brain Interface. Front Neurosci 2018 Dec 13;12:949 [FREE Full text] [doi: 10.3389/fnins.2018.00949] [Medline: 30631262]
7. Kawase T, Sakurada T, Koike Y, Kansaku K. A hybrid BMI-based exoskeleton for paresis: EMG control for assisting arm movements. J Neural Eng 2017 Feb 09;14(1):016015. [doi: 10.1088/1741-2552/aa525f] [Medline: 28068293]
8. Maksimenko VA, Runnova AE, Zhuravlev MO, Makarov VV, Nedayvozov V, Grubov VV, et al. Visual perception affected by motivation and alertness controlled by a noninvasive brain-computer interface. PLoS ONE 2017 Dec 21;12(12):e0188700. [doi: 10.1371/journal.pone.0188700]
9. Musk E, Neuralink. An Integrated Brain-Machine Interface Platform With Thousands of Channels. J Med Internet Res 2019 Oct;21(10) [FREE Full text] [doi: 10.2196/16194] [Medline: 31642810]
10. Zander TO, Kothe C. Towards passive brain-computer interfaces: applying brain-computer interface technology to human-machine systems in general. J Neural Eng 2011 Apr 24;8(2):025005. [doi: 10.1088/1741-2560/8/2/025005] [Medline: 21436512]
11. Maksimenko VA, van Heukelum S, Makarov V, Kelderhuis J, Lüttjohann A, Koronovskii A, et al. Absence Seizure Control by a Brain Computer Interface. Sci Rep 2017 May 29;7(1):2487 [FREE Full text] [doi: 10.1038/s41598-017-02626-y] [Medline: 28555070]
12. Pisarchik AN, Grubov VV, Maksimenko VA, Lüttjohann A, Frolov NS, Marqués-Pascual C, et al. Extreme events in epileptic EEG of rodents after ischemic stroke. Eur Phys J Spec Top 2018 Oct 19;227(7-9):921-932. [doi: 10.1140/epjst/e2018-800019-1]
13. Frolov NS, Grubov VV, Maksimenko VA, Lüttjohann A, Makarov VV, Pavlov AN, et al. Statistical Properties and Predictability of Extreme Epileptic Events. Sci Rep 2019 May 10;9(1):7243. [doi: 10.1038/s41598-019-43619-3]

14. Fisher RS, Krauss GL, Ramsay E, Laxer K, Gates J. Assessment of vagus nerve stimulation for epilepsy: Report of the Therapeutics and Technology Assessment Subcommittee of the American Academy of Neurology. Neurology 1997 Jul 01;49(1):293-297. [doi: 10.1212/wnl.49.1.293]
15. Hodaie M, Wennberg R, Dostrovsky J, Lozano A. Chronic Anterior Thalamus Stimulation for Intractable Epilepsy. Epilepsia 2002 Jun 14;43(6):603-608. [doi: 10.1046/j.1528-1157.2002.26001.x]
16. Kossoff EH, Ritzl EK, Politsky JM, Murro AM, Smith JR, Duckrow RB, et al. Effect of an External Responsive Neurostimulator on Seizures and Electrographic Discharges during Subdural Electrode Monitoring. Epilepsia 2004 Dec;45(12):1560-1567. [doi: 10.1111/j.0013-9580.2004.26104.x]
17. Ovchinnikov A, Lüttjohann A, Hramov A, van Luijtelaar G. An algorithm for real-time detection of spike-wave discharges in rodents. J Neurosci Methods 2010 Dec 15;194(1):172-178. [doi: 10.1016/j.jneumeth.2010.09.017] [Medline: 20933003]
18. van Luijtelaar G, Lüttjohann A, Makarov VV, Maksimenko VA, Koronovskii AA, Hramov AE. Methods of automated absence seizure detection, interference by stimulation, and possibilities for prediction in genetic absence models. J Neurosci Methods 2016 Feb 15;260:144-158. [doi: 10.1016/j.jneumeth.2015.07.010] [Medline: 26213219]
19. Wolpaw JR, McFarland DJ. Control of a two-dimensional movement signal by a noninvasive brain-computer interface in humans. Proc Natl Acad Sci U S A 2004 Dec 21;101(51):17849-17854 [FREE Full text] [doi: 10.1073/pnas.0403504101] [Medline: 15585584]
20. Velliste M, Perel S, Spalding MC, Whitford AS, Schwartz AB. Cortical control of a prosthetic arm for self-feeding. Nature 2008 May 28;453(7198):1098-1101. [doi: 10.1038/nature06996]
21. Fitzsimmons NA, Lebedev MA, Peikon ID, Nicolelis MAL. Extracting kinematic parameters for monkey bipedal walking from cortical neuronal ensemble activity. Front Integr Neurosci 2009;3:3 [FREE Full text] [doi: 10.3389/neuro.07.003.2009] [Medline: 19404411]
22. Ifft PJ, Shokur S, Li Z, Lebedev MA, Nicolelis MAL. A brain-machine interface enables bimanual arm movements in monkeys. Sci Transl Med 2013 Nov 06;5(210):210ra154-210ra154 [FREE Full text] [doi: 10.1126/scitranslmed.3006159] [Medline: 24197735]
23. Hochberg LR, Bacher D, Jarosiewicz B, Masse NY, Simeral JD, Vogel J, et al. Reach and grasp by people with tetraplegia using a neurally controlled robotic arm. Nature 2012 May 16;485(7398):372-375 [FREE Full text] [doi: 10.1038/nature11076] [Medline: 22596161]
24. Mishchenko MA, Gerasimova SA, Lebedeva AV, Lepekhina LS, Pisarchik AN, Kazantsev VB. Optoelectronic system for brain neuronal network stimulation. PLoS ONE 2018 Jun 1;13(6):e0198396. [doi: 10.1371/journal.pone.0198396]
25. Raspopovic S, Capogrosso M, Petrini FM, Bonizzato M, Rigosa J, Di Pino G, et al. Restoring natural sensory feedback in real-time bidirectional hand prostheses. Sci Transl Med 2014 Feb 05;6(222):222ra19-222ra19 [FREE Full text] [doi: 10.1126/scitranslmed.3006820] [Medline: 24500407]
26. Stacey WC, Litt B. Technology Insight: neuroengineering and epilepsy—designing devices for seizure control. Nat Rev Neurol 2008 Feb 26;4(4):190-201. [doi: 10.1038/ncpneuro0750]
27. Bolus MF, Willats AA, Whitmire CJ, Rozell CJ, Stanley GB. Design strategies for dynamic closed-loop optogenetic neurocontrol. J Neural Eng 2018 Jan 25;15(2):026011. [doi: 10.1088/1741-2552/aaa506]
28. Newman JP, Fong M, Millard D, Whitmire C, Stanley G, Potter S. Optogenetic feedback control of neural activity. eLife 2015;4:e07192. [doi: 10.7554/elife.07192]
29. Iseri E, Kuzum D. Implantable optoelectronic probes for in vivo optogenetics. J Neural Eng 2017 Jun;14(3):031001. [doi: 10.1088/1741-2552/aa60b3] [Medline: 28198703]

Edited by G Eysenbach; submitted 22.09.19; peer-reviewed by S Saeb, J Salisbury; comments to author 12.10.19; revised version received 16.10.19; accepted 20.10.19; published 31.10.19.

Please cite as:
Pisarchik AN, Maksimenko VA, Hramov AE
From Novel Technology to Novel Applications: Comment on "An Integrated Brain-Machine Interface Platform With Thousands of Channels" by Elon Musk and Neuralink
J Med Internet Res 2019;21(10):e16356
URL: http://www.jmir.org/2019/10/e16356/
doi:10.2196/16356
PMID:31674923

Original Paper

Towards a Stakeholder-Oriented Blockchain-Based Architecture for Electronic Health Records: Design Science Research Study

Jan Heinrich Beinke[1*], MSc; Christian Fitte[1*], MSc; Frank Teuteberg[1], Prof Dr

Accounting and Information Systems, University of Osnabrueck, Osnabrueck, Germany

*these authors contributed equally

Corresponding Author:
Jan Heinrich Beinke, MSc
Accounting and Information Systems
University of Osnabrueck
Katharinenstraße 1
Osnabrueck, 49069
Germany
Phone: 49 541 969 6185
Email: jan.beinke@uni-osnabrueck.de

Abstract

Background: Data security issues still constitute the main reason for the sluggish dissemination of electronic health records (EHRs). Given that blockchain technology offers the possibility to verify transactions through a decentralized network, it may serve as a solution to secure health-related data. Therefore, we have identified stakeholder-specific requirements and propose a blockchain-based architecture for EHRs, while referring to the already existing scientific discussions on the potential of blockchain for use in EHRs.

Objective: This study aimed to introduce blockchain technology for EHRs, based on identifying stakeholders and systematically eliciting their requirements, and to discuss the key benefits (KBs) and key challenges (KCs) of blockchain technology in the context of EHRs.

Methods: The blockchain-based architecture was developed in the framework of the design science research paradigm. The requirements were identified using a structured literature review and interviews with nine health care experts. Subsequently, the proposed architecture was evaluated using 4 workshops with 15 participants.

Results: We identified three major EHR stakeholder groups and 34 respective requirements. On this basis, we developed a five-layer architecture. The subsequent evaluation of the artifact was followed by the discussion of 12 KBs and 12 KCs of a blockchain-based architecture for EHRs. To address the KCs, we derived five recommendations for action for science and practice.

Conclusions: Our findings indicate that blockchain technology offers considerable potential to advance EHRs. Improvements to currently available EHR solutions are expected, for instance, in the areas of data security, traceability, and automation by smart contracts. Future research could examine the patient's acceptance of blockchain-based EHRs and cost-benefit analyses.

(J Med Internet Res 2019;21(10):e13585) doi:10.2196/13585

KEYWORDS

blockchain; electronic health records; data security; information storage and retrieval

Introduction

In the course of digitization, electronic health records (EHRs) have become one of the most important topics within the health care sector, as they are expected to significantly improve intersectoral collaboration and reduce health care expenses [1]. Given the fact that in many countries, including Germany, there is no government-regulated EHR system, the number of private providers on the market is rapidly increasing. However, for privacy and security concerns, many patients refrain from using an EHR because they fear that the private provider may sell their health data to make profit [2].

Similar to the health care industry, the financial sector also exchanges highly sensitive customer data. In this context, blockchain technology has recently gained attention as a possible solution to secure sensitive transactions. Blockchain technology offers potential, for instance, in the areas of disintermediation, decentralization, the reduction of necessary trust between

business partners, improved protection against data manipulation, and increased automation through smart contracts [3-5]. This leads to the question, whether blockchain technology could also be able to ensure the use of EHRs. Although some authors have already analyzed the general potential of blockchain for the health care sector, a specific suggestion on how this potential can be achieved is currently missing [6]. Kuo et al, for instance, investigate the application possibilities in the field of biomedical and health care applications [7]. Although this study provides an interesting overview of the potential and challenges of blockchain-based applications, there is a lack of specific possibilities for implementation. In detail, a systematic requirement analysis for the respective field of application is missing. Further contributions present possible blockchain-based architectures for EHRs [8-10]. However, in these contributions, it often remains unclear on which scientific or practical basis the proposed architecture has been developed. Moreover, the requirements and interests of the stakeholder in the health care system are often not taken into account. In contrast to existing blockchain architectures for EHRs, our approach follows a structured scientific development and aims to include the stakeholders' perspectives. Furthermore, there are contributions that identify the possible advantages and disadvantages of blockchain-based EHRs for different actors [6,7]. However, according to the authors, these are not complete, as not all relevant actors were investigated. Consequently, it can be stated that, to date, there is no contribution that presents a blockchain-based EHR architecture that is based on (1) multimethodically and (2) systematically collected requirements for (3) all relevant stakeholders in the health sector and also elaborates (4) the key benefits (KBs) and (5) key challenges (KCs) of the developed architecture. Similar demands for future research in this context are confirmed by other authors [7,11,12]. This motivates us to investigate the following research questions (RQs):

> *RQ 1: Which stakeholders have an interest in EHRs and what are their specific requirements for an EHR?*
>
> *RQ 2: How can these requirements be implemented in a blockchain-based architecture?*
>
> *RQ 3: Which key benefits and key challenges does the proposed blockchain-based EHR architecture provide?*

The paper is structured as follows: The methodological framework is presented in the Methods section. On the basis of the design science research (DSR) method, we first identified stakeholders and collected their respective requirements using a systematic literature review and 9 interviews of health care experts. The subsequent evaluation cycle was carried out through 4 workshops with 15 participants in total, for example, health care professionals, information systems experts, and lawyers. The Results section presents the consolidated requirements of the stakeholders, the architecture, which was developed based on the identified requirements, and results from the evaluation cycles. The Discussion section focuses on the KBs and KCs that have been derived from the literature, the expert interviews, and the workshop discussions. Solutions for the identified KCs were proposed in the form of 5 recommendations for action. Limitations of the study have been elaborated, and possible further research perspectives have been pointed out. Our contribution contains interesting findings for science and practice alike, especially for EHR providers. The systematic requirement analysis can be used as a basis for the (further) development of other architectures. Besides, we have shed some light on the effects of the digital transformation on the health care sector. Another benefit of the developed architecture is that it can be prototypically implemented by companies and thus tested in a real context. The insights gained could serve for refinement of the architecture, which again allows for new findings with a different focus, such as acceptance investigations or a cost-benefit analysis.

Methods

To answer the identified RQs, we applied the DSR paradigm that addresses human-relevant problems using innovative artifacts and simultaneously contributes new knowledge for the scientific community [13]. Designed artifacts are not only useful but also crucial to understand the problem itself. Thus, DSR is especially suitable to serve as the basis for the development of a blockchain-based EHR architecture. Although we have addressed the major security issues that hamper the diffusion of EHRs, we have involved the stakeholders in the development phase, which fosters their awareness and understanding of the technology.

The development of a solution is influenced by the environment and the existing body of knowledge as visualized in Figure 1 [14]. Thereby, the relevance of the identified problems represents the connection with the environment, whereas the link to the knowledge base is represented by the recognition of results of relevant works and a rigorous application of research methods [15]. The design cycle within the DSR is altering between development and evaluation. In this study, we applied a literature review and 9 qualitative interviews to elaborate the stakeholders' requirements. On the basis of these results, we developed a concept for a blockchain-based architecture. The intermediate results were evaluated in 4 workshops with 15 health care experts.

To identify related work, we conducted a structured literature analysis according to vom Brocke et al [16]. The search term (Blockchain OR distributed ledger) AND (EHR OR "Electronic Health Record" OR "EPR" OR "Electronic Patient Record" OR PHR OR "Personal Health Record" OR EPA OR EGA OR (Elektronische OR Digitale" AND Patientenakte OR Gesundheitsakte)) was applied to the information systems and health care databases such as EBSCOhost, Emerald, IEEE Xplore, MEDLINE, ProQuest, PubMed, ScienceDirect, Scopus, Wiley, and Google Scholar.

On the basis of the outcome of our literature search, we identified all stakeholders who probably use EHRs. Although EHR providers are also key stakeholders, we excluded them from our investigation because their interest rather lies in the requirements of their customers. The remaining stakeholders can be categorized into 3 groups (Figure 2). Primary stakeholders are directly involved in providing health care, for example, physicians, caregivers and nurses, therapists, pharmacists, clinics and hospitals, laboratories, care services,

nursing homes, and the patient themselves [8,17,18]. The group of secondary stakeholders includes insurances, family and relatives, and employers, whereas the tertiary stakeholder group comprises society, research institutes, public authorities, and the health care industry.

In the next phase, we collected the respective stakeholders' requirements for EHRs from the relevant literature. To enrich the scientific perspective, we additionally conducted interviews with 9 health care experts (Table 1), who represent the mentioned stakeholder groups in the period from July to October 2018 [19]. The interviews were recorded and analyzed individually by the authors. In the Results section, we have consolidated the respective requirements. By means of the outcomes of the 4 subsequent workshops with 15 participants in total, for example, health care professionals, information system experts, and lawyers, we are in a position to evaluate our presented concept [19].

Figure 1. Design science research method for concept development (Hevner et al). eHealth: electronic health; EHR: electronic health record.

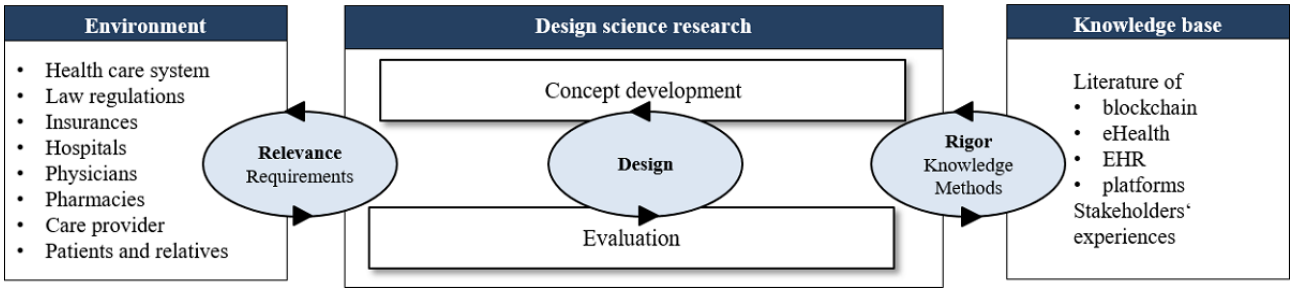


Figure 2. Overview of stakeholder groups.

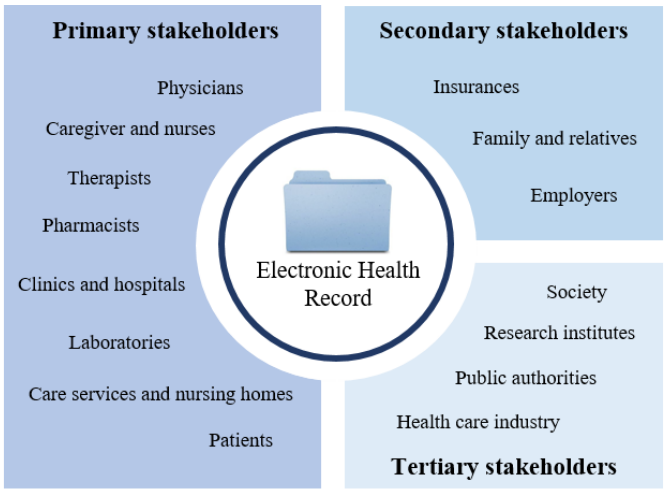


Table 1. Overview of interviewed experts.

No	Description	Work experience (years)	Duration (min)
E1	Pharmacist	22	40
E2	Health care consultant	2	31
E3	Pharmacist	18	27
E4	Founder and developer of an electronic health app	1	22
E5	Male nurse and case manager	30	24
E6	Health care consultant	7	36
E7	Managing director of an electronic health record	7	24
E8	Nurse	9	22
E9	Physician	10	32

Results

Requirements

On the basis of our systematic literature review and the 9 expert interviews, we identified a total of 34 requirements (R) for blockchain-based EHRs (Table 2). The requirements were assigned to the 3 stakeholder groups.

Table 2. Consolidated requirements of stakeholder groups for a blockchain-based electronic health record.

No and group	Requirement	References	E1	E2	E3	E4	E5	E6	E7	E8	E9
Primary stakeholders											
R1	Data security[a]	[7,12,20]	__[b]	x	x	x	x	x	x	—	x
R2	Data privacy[a]	[7,12,20]	x	x	x	x	x	x	x	—	x
R3	Access/permission control, data sovereignty	[8,20-22]	x	x	—	x	—	—	x	—	x
R4	Identity confirmation[a]	[8,20]	x	x	—	x	—	—	x	—	x
R5	Manipulation protection/data integrity[a]	[7,21,23,24]	x	—	x	x	—	x	x	—	x
R6	Complete health record[a]	[21]	x	x	x	x	x	—	x	x	x
R7	Performance[a]	[7]	—	x	—	—	x	—	x	—	x
R8	User friendly design[a]	[25]	x	x	x	x	—	—	x	x	x
R9	Context-specific information[a]	[25]	—	x	x	—	x	x	x	x	—
R10	Data and file storing[a]	[20]	x	x	x	x	x	x	x	—	x
R11	Data and file sharing[a]	[8,12,20,21]	x	x	x	x	x	x	x	x	x
R12	Interoperable and consistent data standards[a]	[12,20]	x	x	—	x	—	x	x	x	x
R13	Intersectoral communication[a]	[12,21]	x	x	x	—	x	x	x	—	x
R14	Ensuring trusted relationships[a]	[12,20]	—	x	—	—	—	—	x	—	
R15	CRUD[c] rights	[8,21]	x	—	x	—	x	—	x	x	x
R16	Verification modus	[8,21]	x	x	—	—	—	—	x	—	x
R17	Emergency pass	[26]	—	x	x	—	—	—	x	—	x
R18	Medication/care plan	[27,28]	x	—	x	x	—	x	—	—	x
R19	Tracking of state transitions[a]	[8]	x	—	x	—	x	x	x	x	x
R20	General administrative issues	—	—	x	—	—	x	—	x	—	—
R21	Synchronization to off-chain data[a]	[8]	—	—	—	—	—	—	—	—	—
R22	Notification services	[8]	x	x	—	x	x	—	x	—	x
R23	Modularity[a]	[20]	—	—	—	x	—	—	x	—	—
R24	Patient centration	[12]	x	—	—	—	x	—	—	—	x
R25	Workflow support[a]	[29]	x	x	—	x	—	x	x	—	—
R26	Integration into existing systems[a]	[8,29]	x	x	x	x	—	x	—	x	x
R27	Transfer sheet	—	—	x	x	—	x	x	x	—	x
Secondary stakeholders											
R28	Scalability[a]	[7,12,20]	—	—	—	x	—	x	x	—	—
R29	Invoice management	—	x	x	x	—	—	—	x	—	—
R30	Modus for relatives	—	—	x	—	—	x	—	x	—	—x
Tertiary stakeholders											
R31	Access to consolidated data[a]	[10]	—	x	—	—	—	—	x	—	x
R32	Statistics	[10,30]	—	x	—	x	—	—	x	—	—
R33	Clinical research	[31]	—	x	—	x	—	—	x	—	x
R34	Predictive analyses	[31]	—	x	—	x	—	—	x	—	x

[a]To avoid multiple entries, requirements that apply to all stakeholder groups, for example, data security and data privacy, are listed once.

[b]Not applicable.

[c]CRUD: create, read, update, and delete.

Primary Stakeholders

The most important EHR concerns of this stakeholder group are connected to data security and data privacy (R1 and R2) [7,12,20]. Among other things, because of its distributed structure, the use of consensus mechanisms, and cryptographic methods, blockchain technology offers a high potential to counteract these concerns [8]. To protect the highly sensitive health data, patients should have full access and permission control and the possibility to precisely designate each actor involved (eg, physicians and relatives) [8,20,21]. Furthermore, it is essential that they retain data sovereignty, which means that the decision who has access to what data are incumbent on them (R3 and R4) [22]. For example, if patients seek a second medical opinion, they might want to avoid that the first diagnosis affects the second physician.

Nevertheless, an EHR should contain the patients' complete treatment history to provide involved physicians with the full picture and allow for optimal treatment (R6) [21]. Thereby, it is of utmost importance that the relevant information are quickly accessible [7] but cannot be manipulated [7]. It must be ensured that these data can be updated but not manipulated [7,8,21,23]. Moreover, all stakeholder groups demand for a user-friendly design and context-specific information to avoid an information overload [25].

To enable intersectoral communication, storing, retrieving, and sharing of files and data turned out to be of high relevance (R10 and R11) [8,12,20,21]. These files and data formats should comply with consistent standards (R12) [12,20].

Furthermore, blockchain technology is capable of reducing the necessary trust between the involved actors because transactions support the aforementioned mechanisms (R14) [12,20]. In addition, blockchain technology supports data integrity, as each transaction is recorded (R5) [7]. Besides, data protection is of particular importance (R2). This can basically be supported by blockchain technology, but limitations have to be stated here. When analyzing the metadata, conclusions could be drawn about single individuals under certain circumstances [8,32]. For example, long-term monitoring of common diseases and frequently visited health care actors can provide systematic data analyses and allow for useful conclusions. However, as this conflicts with the goals of data protection, the implementation of anonymization mechanisms, such as those currently used by some cryptocurrencies as Zcash, is a suitable option for anonymizing transaction data [33].

Within this context, access control (R3) and identity confirmation (R4) again play a special role. Both could be carried out by the state, as is already the case in current health systems, such as in Estonia. Furthermore, the allocation of create, read, update, and delete (CRUD) rights must be mentioned here [8,21]. In this context, it is indispensable that specific organizational units are granted certain rights, for example, the right to insert new documents into the EHR. However, the performance of the application (R7) must be guaranteed for all transactions carried out to be attractive for the user groups [7]. The allocation or modification of, for instance, viewership rights should be tracked accordingly (R19) [8]. Before the release of data that have been newly inserted by health care professionals, the content could be checked by the respective patient by means of a *verification mode* (R16) [8,21].

Particularly, the integration of existing systems (R26; eg, hospital information system, pharmaceutical medication plan, and physicians' patient administration system) is of importance, as this would require fewer adjustments by all actors involved [8,29]. Equally, the integration of existing workflows is requested (R25) [29]. In addition, general administrative issues should be covered, for example, the status of sick leave or the current insurance status. (R20). A continuous further development and the associated expansion of the functional scope are also planned and can be implemented in the form of individual modules (R23) [12].

Another functionality that can improve the intersectoral communication is a transfer sheet that contains all relevant information necessary for a patient's transfer from one institution to another, for example, from a hospital in a nursing home (R27).

These transfer sheets include information on previous treatments and convey instructions to the next health care actor in charge.

In this context, there are features, such as the emergency pass, that provide all relevant emergency data (R17; eg, blood group and allergies) and a medication plan, which provides details on dosage, side effects, and drug interactions of the medications to be taken (R18) [26-28]. Notification services (R22: eg, vaccination plans) and context-specific information (R9), for example, alarm triggering when permissible vital parameters are exceeded, represent further useful enhancements [8].

Interoperability and consistent data standards (R12) as well as automation through smart contracts can also improve intersectoral communication (R13) [12,21]. However, there is often the problem that data are available in various formats and at different storage locations within the health care system, which considerably complicates a smooth communication. To enable a preferably intuitive handling for the patients, the user interface should be designed as user friendly (R8) and patient centric as possible [12,25].

Secondary Stakeholders

To ensure the greatest possible support from all relevant actors, the requirements set by the secondary stakeholders also have to be taken into account. For instance, it is necessary for EHR solutions to be scalable, as insurances will probably provide these to all their policyholders. Furthermore, a *relatives mode* would provide a significant added value for patients who cannot maintain files themselves (R30). In addition, health insurance companies could use the system to manage their prescriptions and monitor the compliance and efficacy of therapies.

Tertiary Stakeholders

For tertiary stakeholders, interesting possibilities arise from the analysis of anonymized consolidated data (R31) [10]. These data could be useful for statistical analyses about specific diseases, the efficiency of therapies (R32), and clinical research (R33) [31]. Research institutes and governments could use the data to predict diseases such as flu outbreaks (R34) [31]. However, this requires that the data set is as complete and consolidated as possible. It is, for example, also conceivable to monetize (anonymized) data. This means that institutes that seek information would have to pay the respective patients for the permission to use specific data. However, data can only be passed on actively by the user (R3, R11, and R24), and the respective institution has to clearly specify the intended use. In case of infringement, effective penalties (eg, high fines, imprisonment, or exclusion from network) could be imposed depending on the severity of the breach.

Architecture

As part of the design science approach (first iteration), we developed an initial concept for blockchain-based EHRs within a workshop on the basis of the multimethodically collected requirements analysis. First, the identified requirements have been incorporated into the development of the concept (Figure 3).

Ølnes describes the (bitcoin) blockchain as an information infrastructure that can always be developed further [34]. In his argumentation, he refers to the definition of Hanseth and Lyytinen [35], who define the concept of information infrastructure as a common, open (and unlimited), heterogeneous, and evolving sociotechnical system consisting of a set of information technology (IT) skills and their users as well as operations and design communities. Blockchain technology offers a multitude of possible variations, for example, adding or removing actors or smart contracts (R3), using the implemented consensus mechanism. This enables that the system can be customized to meet specific needs and requirements. The developed concept (Figure 3) takes into account 3 basic options of data management: (1) it offers the possibility to interlink or reference already existing data sources; (2) it provides the possibility to store information encrypted in the blockchain; and (3) it allows to store and reference data from different data sources encrypted in the blockchain (R10 and R11). We deliberately avoided a uniform procedure to guarantee access to as many people as possible and have the highest possible flexibility. People living in countries with an appropriate digital health infrastructure are likely to choose option 1 or option 2 (R21). Although the latter offers the advantage that the data are always available, which reduces dependency on other actors, the disadvantage is the direct (encrypted) storage on the blockchain. Owing to the data protection requirements (eg, General Data Protection Regulation) and technical advantages, the authors recommend referencing the data and using the blockchain as an access management solution. This also makes it easier to adequately manage the large amount of data that are generated, for instance, during clinical studies. Patients might be notified about data changes, which they would have to approve or reject (R16).

All processes within the EHRs should be documented on the blockchain to be able to ensure complete traceability in the event of data breaches or legal disputes (R19). In a working paper for the United Research Institute for Social Development, Scott points out that blockchain-based currencies, so-called cryptocurrencies (eg, bitcoin), are a relatively simple way of managing cash holdings in countries with a weakly developed financial infrastructure and safely handling payment transactions on the spot in insecure, informal environments [36]. Similarly, the concept offers the possibility of independently maintaining the data (option 2), which is especially valuable for people who live in countries with poor (digital) health infrastructure. The access to the blockchain should be realized through a registration with an appropriately certified institution to prevent aggregation of irrelevant or trivial data (R3). Similar considerations can be found in a study by Ekblaw et al [8]. Furthermore, institutions can gain access to aggregated, anonymized data if they provide computer resources through their activities as Miner and thus ensure the network's trustworthiness (R31). This incentive system is based on a study by Ekblaw et al [8]. The functionalities such as the access rights of the other actors to the patient data are defined by smart contracts. Automated contracts between the other players are also possible (R13, R14, and R12). In this way, the entire process can be supported, from admission and diagnosis to treatment and any associated rehabilitation to discharge and final consultation (R25).

After presenting the first concept in Figure 3 and the discussion with health care experts about the data sources and functionalities, we substantiated the concept into an n-tier architecture. The blockchain-based architecture can be divided into 5 layers as shown in Figure 4: data layer at the bottom, data access layer, application logic layer, application layer, and presentation layer. In all layers, ethical and legal implications have to be considered.

Figure 3. Electronic health record concept blueprint. Rehab: rehabilitation.

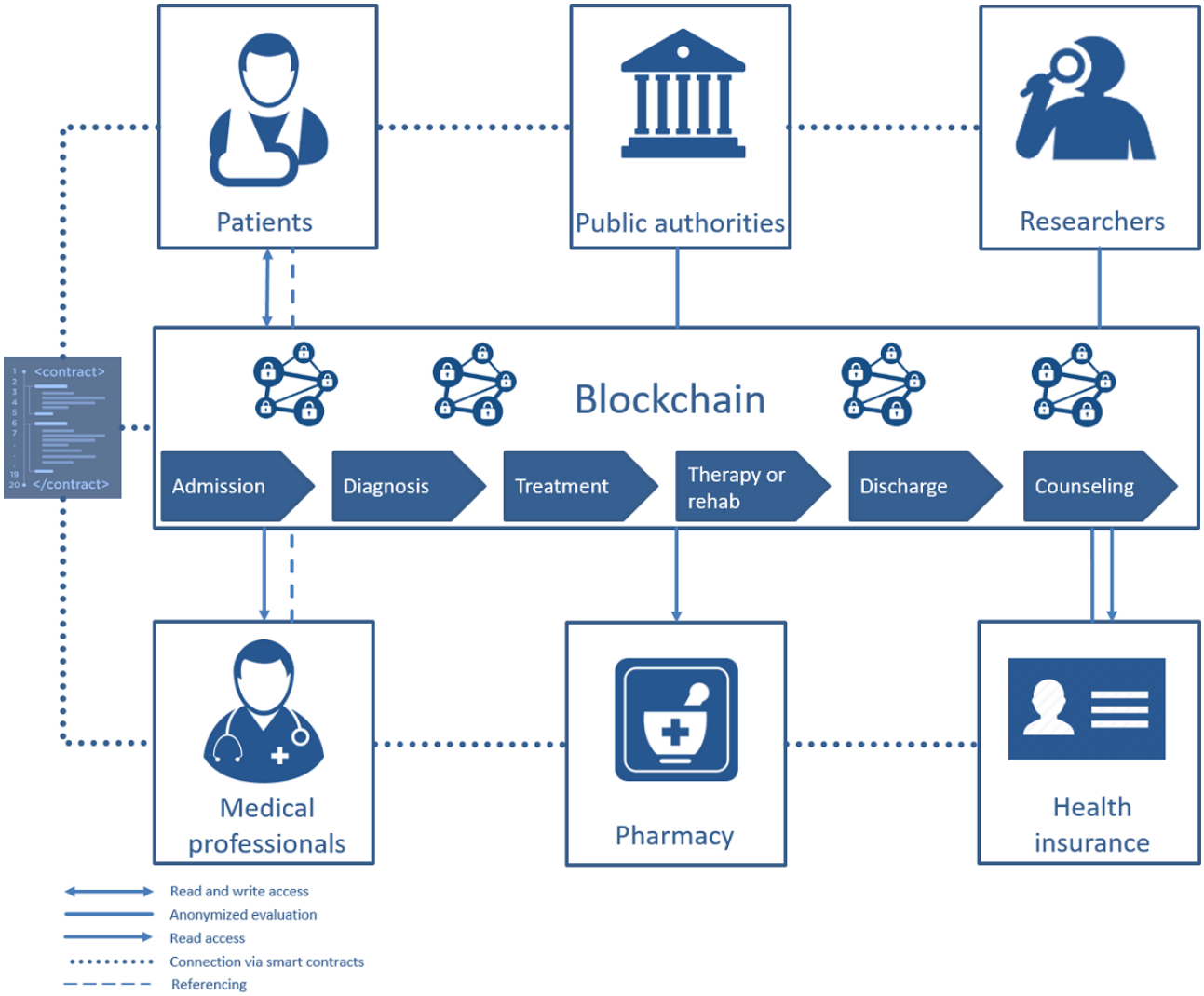
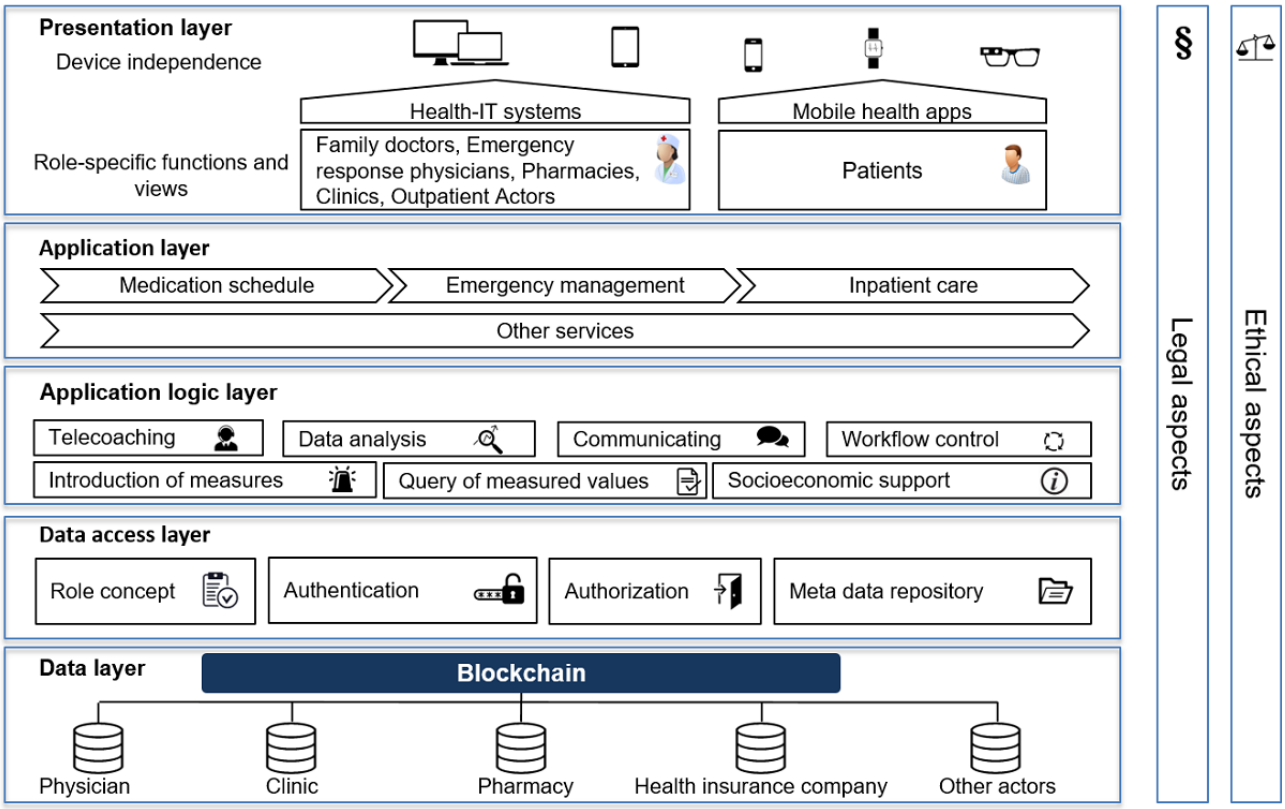


Figure 4. Five-tier architecture. IT: information technology.



Data Layer

The data are standardized to ensure compatibility (R12) and intersectoral communication (R13). The integration platform provides defined communication patterns and interfaces for structured information and document exchange on the Fast Healthcare Interoperability Resources standard, the standard profiles of the Integrating the Healthcare Enterprise, the HL7 family (Health Level 7), openEHR, and the xDT family. Conformity with International Standard Organization (ISO) 21090 (health informatics — harmonized data types for information interchange) and ISO 13606 (health informatics — EHR communication) must be taken into account. For the generic description of the respective information unit, the semantics of the respective useful and information elements used are stored in the integrated metadata repository. The respective treating actor in the health sector has the possibility to provide data for the respective patient and document changes (R15 and R19).

Data Access Layer

Data from different sources are linked to the blockchain to ensure that the patient's health records are as complete as possible (R8). Access control is important to ensure that the respective actors only have access to the relevant and released data (R3 and R2). In this context, the implementation of a role concept is of great importance. In addition, a precisely designed role concept in conjunction with a regulated access control renders data manipulation (R5) more difficult. Besides, a meta-data repository can support the integration and organization of relevant metadata (R26). For example, the data from the disparate systems can be better related to each other, and discrepancies, gaps, and metrics can be identified and addressed at the data structure level.

Application Logic Layer

The application logic layer controls the workflow of the different applications and systems (R25). Thus, data are prepared and made available for different purposes (R9 and R31). This is particularly interesting for the research area and enables various statistical evaluations (R32, R33, and R34).

Application Layer

On the basis of the connected data (R23), various services can be set up. A modular concept allows each stakeholder to individually configure their personal dashboards with only relevant data (R9). According to the experts surveyed, this leads, in particular, to added values for emergency passes (R17), medication plans (R18), and transfer sheets (R27). At this point, for example, the invoice management (R28) of health insurance companies can be applied. Furthermore, the opportunity to delegate administration rights to relatives could also be addressed in the form of a *relatives mode* (R30).

Presentation Layer

The presentation layer describes how the services are displayed to the end user. The range of functions for a user depends on his role, for example, more comprehensive functions and data are available for emergency physicians than for pharmacies. It is particularly important that the platform can be accessed either using responsive Web applications or native applications on almost all end devices and operating systems (R8). Modern smartphones, in particular, offer interesting functions with additional security measures such as fingerprint sensors and iris

scanners (R1 and R3). Thus, information can also be prepared and displayed in a context-specific manner (R9). For example, different information can be displayed for a nurse wearing augmented reality glasses than for a pharmacist who receives additional information about the current medication plan and can thus be informed about drug interactions at an early stage.

Evaluation

In our first evaluation cycle, we discussed our initial categorization of the identified stakeholders into 3 stakeholder groups (Figure 2) with 3 health care professionals and 3 information systems experts. As this categorization turned out to be too rough for the design of the EHR architecture, we refined it by splitting the stakeholders into patients, medical professionals, public authorities, pharmacies, researchers, and health insurances (Figure 3). In the second evaluation workshop with 2 health care professionals and 3 information system experts, we discussed the databases that should be included in the concept. The participants evaluated the physicians, clinics, pharmacies, and health insurance databases as most important for the EHR architecture (see data layer in Figure 4). The remaining sources were summarized in other actors.

A draft of the first concept (Figure 3) was presented to 5 health care professionals at the third evaluation workshop. Each participant added specific use cases from his profession, which we gradually included in the concept. Owing to the diversity of the use cases, we decided to begin with the 6 aspects of admission, diagnosis, treatment, therapy or rehabilitation, discharge, and counseling. In the final architecture, we reduced the complexity by implementing an application layer that did not include individual use cases but showed exemplary modules of the EHR system.

At the last evaluation workshop, we presented the final architecture (Figure 4) to 4 health care professionals, 2 lawyers, and 3 information system professionals. Additional functionalities and several data standards were discussed and incorporated into the architecture concept. Finally, we elaborated the KBs and KCs of the proposed solution.

Discussion

Principal Findings

After identifying the relevant stakeholders and their 34 respective requirements with the help of a literature review and expert interviews, we developed the first concept for a blockchain-based EHR. The concept was subsequently evaluated with the experts again to build a 5-tier architecture, which is presented in Figure 4. The development and introduction of blockchain-based EHRs are accompanied by several KBs and KCs. In the following, the KBs are presented before the KCs are critically discussed and possible solutions, in the form of 5 recommendations for action, are proposed.

Key Benefits

We identified 12 KBs of a blockchain-based EHR architecture, which we summarized with their respective sources from the literature and expert interviews in Table 3. The decentralization of the blockchain is the first KB (KB1) to be mentioned because distributed systems are usually less susceptible to system failures. Thus, there is no single point of failure (KB2). Moreover, the failure of individual nodes does not have significant effects on system security. A further advantage of the blockchain is that it has implemented various mechanisms (eg, consensus mechanism and cryptographic procedures) that render data manipulation difficult (KB3). The blockchain's relatively high security, ensured among others by the cryptographic algorithms and decentralization, constitutes another benefit (KB4). Furthermore, the blockchain allows the tracking of entries, which makes incorrect treatment decisions traceable (KB5) and increases the patient safety. In addition, it is possible to store the entire treatment history, which significantly increases the scope of information available to the treating actors (KB6) and contributes to a general improvement in treatment quality. Smart contracts can help to automate certain processes (KB7; eg, referrals to specialists, ordering [individualized] medication). In addition, data sovereignty is firmly transferred to the patient (KB8) so that the user is the *master of his own data*. The fact that all relevant data can be made available to the corresponding actors quickly and in a standardized format, significantly increases intersectoral communication (KB9). Furthermore, it is conceivable that service providers process both invoicing and payment transactions directly using the blockchain (KB10). Possible new business models are emerging, for example, through the systematic evaluation of data or brokerage services (KB11). In summary, it can be said that the presented concept offers advantages at various levels, including technical (eg, data security), organizational (eg, intersectoral communication), and economic issues (eg, new business models). All in all, the advantages of blockchain-based EHRs could significantly improve the status quo of patient-oriented treatment (KB12).

Table 3. Key benefits (KB) of a blockchain-based electronic health record architecture.

No	Key benefits	References	Experts
KB1	Decentralization	[7,23]	E7
KB2	No single point of failure/vulnerability	[7]	E7
KB3	Tamper proof	[7,23,24]	E1, E2, E4, E7, E9
KB4	Data security	[23,37]	E2, E4, E7, E9
KB5	Traceability of entries	[7,23]	E1, E2
KB6	Overview of all health-related data	[21]	E7, E9
KB7	Automation by smart contracts	[38]	—[a]
KB8	Data sovereignty for patient	[7,12,23]	E1, E2, E5, E6, E7, E9
KB9	Improved intersectoral collaboration through file and data sharing	[23]	E1, E2, E3, E4, E5, E7, E8, E9
KB10	Integrated payment application	[23]	E2, E7
KB11	New mining business models for data analysis	[8,21]	E7
KB12	Patient-oriented treatment	[12]	E1, E3, E8, E9

[a]Not applicable.

Key Challenges

Despite all advantages, a blockchain-based EHR still faces some major challenges that are summarized with their respective sources from the literature and expert interviews in Table 4. The high energy consumption, primarily because of the use of the proof-of-work consensus mechanism, is often cited as a major challenge associated with the use of blockchain technology (KC1). This leads to high transaction costs (KC2) because of both high energy consumption and the required hardware resources (KC3). However, these challenges can primarily be addressed by 2 measures. First, other consensus mechanisms such as proof-of-stake could significantly reduce the electricity consumption; second, regulated access, in the sense of a consortium blockchain, could keep the required hardware investments manageable. Regarding access regulation, however, the question arises as to which organization is responsible (KC4). We currently consider the state or a consortium consisting of different stakeholders to be potential access regulators. Both could equally be considered when it comes to the questions of responsibility and accountability for (further) development as well as the administration of the system (KC5).

However, the fact that systematic analyses based on metadata offer comprehensive assessment possibilities and allow for conclusions, they have to be seen as a technical challenge at the same time (KC6). Although no satisfactory solution to this problem has yet been found, the cryptocurrency community is currently addressing it, for example, by means of the anonymity mechanism of the cryptocurrency Zcash [33]. Another technical aspect that needs to be addressed in the future is the so-called 51% attacks (KC7). However, they are very unlikely in a consortium. Moreover, the processing speed of blockchains is relatively slow compared with conventional databases (KC8). For example, the Bitcoin blockchain needs up to 10 min to write transactions into a new block. However, it is usually not necessary for these data to be available to other participants in real time. In addition, this time delay only affects the writing, not the reading of transaction data. The verification of imported data, which is mandatory for the highest possible data quality and timeliness (KC9), constitutes another major challenge. In principle, the respective user could confirm the entered data again before it is finally written down. However, this could also result in disadvantageous delays or users rejecting or concealing the doctor's findings. Finally, 2 further challenges can be identified, the primary aim of which is to motivate the involved actors. On the one hand, the added value must be demonstrated to patients and health care professionals, and the necessary (technical) skills in handling such complex applications must be developed (KC10). But the benefits for the remaining stakeholders also have to be demonstrated (KC11). According to the experts, this can best be accomplished by focusing on the expected, significantly more favorable cost structure in the long term, which clearly stands out against the costs of the existing, highly fragmented, and thus relatively cost-intensive IT-system landscape. In particular, automation with smart contracts could also address a high savings potential here.

As blockchain technology is relatively new, standards are lacking, for example, regarding reference architectures and interfaces to other blockchains (KC12).

Table 4. Key challenges (KC) of a blockchain-based electronic health record architecture.

No	Key challenges	References	Experts
KC1	High energy consumption	[39]	—[a]
KC2	High and unpredictable transactions costs	[10,12,20,24,37]	E2, E7
KC3	Requires high storage, bandwidth and computational power, low scalability	[7,10,12,20,39,40]	E7
KC4	Access and authorization issues	[10]	E5, E6, E7
KC5	Accountability for development and administration	[23]	E5, E7
KC6	Public availability of transactions	[7,10,12,20,24,37,39,41]	E2, E3, E7, E9
KC7	51% attack	[7,37]	—
KC8	Slow processing speed	[7,12,39]	E2, E4, E7
KC9	Data imports need verification	[10,20]	E2, E4, E7, E9
KC10	Technical skills of patient and health care professional	[10]	E2, E5, E7, E8, E9
KC11	Incentives for provision of computational resources	[23]	E2, E4, E7
KC12	Standardization	[20]	E1, E3, E5, E6

[a]Not applicable.

To address these challenges, we derived 5 recommendations for action for science and practice.

Recommendation 1

The first recommendation for action is the development of comprehensive standards (KC12). It is conceivable to track and co-design the current standardization attempts such as ISO/TC 307 (blockchain and distributed ledger technologies) and specify this standardization for blockchain-based applications in the health sector. In addition to fundamental topics such as terminology, vulnerabilities, and reference architectures, legal issues such as the legal validity of smart contracts or governance aspects are also of interest. Standardization could thus also help to shape responsibilities for development and administration, especially in the context of IT governance (KC5).

Recommendation 2

The authors recommend the formation of a cross-stakeholder consortium that addresses both technical challenges (KC1, KC6, KC7, and KC11) and organizational challenges (KC4 and KC5). This consortium could establish a consortium blockchain (also called hybrid blockchain) and simplify access controls. Furthermore, consortia offer the advantage of forming a concrete entity capable of acting, which can, for example, improve the representation of interests. In principle, the question arises as to who will be involved in the consortium and how this involvement will look like. To this end, the members of the consortium should be elected from all relevant stakeholder groups.

Recommendation 3

Researchers and companies should continue to work on advancing blockchain technology. Thereby, encryption mechanisms must protect the data as effectively and efficiently as possible (KC6 and KC8), and consensus mechanisms must work as resource saving as possible (KC1 and KC3) without compromising security. Reducing resource use would strengthen environmental sustainability and reduce long-term financial operating costs in the form of transaction costs (KC2).

Recommendation 4

Costs are a decisive factor for every project. As such a project would involve high (financial) costs (KC1, KC2, KC3, and KC11), a further focus should be placed on the development of business models. Perspectives are opened here, for example, in the (anonymized) evaluation of data that can provide interesting insights for the (further) development of drugs, treatments, and therapies.

Recommendation 5

The 5th recommendation for action states that users must be empowered and trained to use current information and communication technologies (ICTs; KC10) and learn more about their current state of health (health literacy) to be able to trace findings to some extent and thus reduce false entries in the system (KC9). The authors are aware that probably not both goals can be realized for every user. Nevertheless, the aim should be to inform the user as well as possible about the possibilities of modern ICT and their state of health. At this point, concepts such as digital nurses or digital learning platforms could offer interesting perspectives.

Limitations

Although the presented blockchain architecture is supposed to take all stakeholders' requirements into account to provide optimal conditions for a qualitative health care supply, it has its limitations. First, we have not been able to interview representatives of each identified stakeholder group. For example, the specific requirements of insurance companies or research institutes need to be examined more closely. Furthermore, we did not include the opinion of the most important stakeholders in the adaption of EHR, namely the patients. The personal attitude toward technological innovations is highly subjective and depends on age, technological affinity, and pre-experiences. Therefore, quantitative surveys should be conducted to investigate the acceptance of blockchain-based EHRs. The fact that the evaluation of our results will require

deep technical and organizational know-how about blockchain technology constitutes another important challenge.

Conclusions

The aim of our analysis was to investigate whether and how blockchain technology can be used for EHRs. To do so, we applied the DSR paradigm. First, we identified 15 stakeholders and categorized them into 3 groups. With the help of a structured literature review and 9 expert interviews, we collected 34 specific requirements for EHRs (RQ1). In the next phase, we drafted the first concept for a blockchain-based architecture. Within 4 iterative evaluation cycles, we developed a 5-tier architecture that takes the identified requirements into account (RQ2). Finally, we discussed KBs and KCs of our proposed solution (RQ3) and derived 5 recommendations for action to address the KCs.

We conclude that blockchain technology offers considerable potential to improve EHRs. In contrast to currently available EHR solutions, blockchain technology offers improvements, for instance, regarding data security, traceability, and automation by smart contracts. We identified 12 KBs, which can be achieved by using blockchain technology for EHRs.

Nevertheless, there are still some KCs that need to be overcome (Table 4). Future research should address ethical, social, environmental, technological, and economic implications. First of all, research potential can be identified in the investigation of incentive programs for providing computational resources. This is connected to the question which business model would be most suitable to offer a blockchain-based EHR. In this context, cost-benefit analyses should be conducted. Furthermore, issues according to data security/privacy and the attack threat need to be analyzed. Finally, it is crucial that especially patients accept the technology. As discussed above, quantitative acceptance investigations are needed to improve applications and enhance dissemination. We expect new blockchain-based applications to emerge in the health care sector that have the potential to substantially improve existing solutions and thus the quality of health care supply.

Acknowledgments

The authors would like to thank Ms Imhorst and the reviewers for their constructive and substantial feedback.

The authors acknowledge the support by Deutsche Forschungsgemeinschaft (DFG) and Open Access Publishing Fund of Osnabrück University.

Conflicts of Interest

None declared.

References

1. Häyrinen K, Saranto K, Nykänen P. Definition, structure, content, use and impacts of electronic health records: a review of the research literature. Int J Med Inform 2008 May;77(5):291-304. [doi: 10.1016/j.ijmedinf.2007.09.001] [Medline: 17951106]
2. Tang PC, Ash JS, Bates DW, Overhage JM, Sands DZ. Personal health records: definitions, benefits, and strategies for overcoming barriers to adoption. J Am Med Inform Assoc 2006;13(2):121-126 [FREE Full text] [doi: 10.1197/jamia.M2025] [Medline: 16357345]
3. Avital M, Beck R, King J, Rossi M, Teigland R. Jumping on the Blockchain Bandwagon: Lessons of the Past and Outlook to the Future. In: Proceedings of the International Conference on Information Systems. 2016 Presented at: ICIS'16; December 11-14, 2016; Dublin, Ireland.
4. Rückeshäuser N. Do We Really Want Blockchain-Based Accounting? Decentralized Consensus as Enabler of Management Override of Internal Controls. In: Proceedings of the 13th International Conference on Wirtschaftsinformatik. 2017 Presented at: WI'17; February 12-15, 2017; St. Gallen, Switzerland.
5. Weber I, Xu X, Riveret R, Governatori G, Ponomarev A, Mendling J. Untrusted Business Process Monitoring and Execution Using Blockchain. In: Proceedings of the International Conference on Business Process Management. 2016 Presented at: BPM'16; September 18-22, 2016; Rio de Janeiro, Brazil p. 329-347. [doi: 10.1007/978-3-319-45348-4_19]
6. Linn LA, Koo MB. Blockchain for Health Data and Its Potential Use in Health IT and Health Care Related Research. In: Proceedings of the OSEHRA Innovation Webinar Series. 2016 Presented at: OSEHRA'16; February 20, 2016; Gaithersburg, Maryland.
7. Kuo TT, Kim HE, Ohno-Machado L. Blockchain distributed ledger technologies for biomedical and health care applications. J Am Med Inform Assoc 2017 Nov 1;24(6):1211-1220 [FREE Full text] [doi: 10.1093/jamia/ocx068] [Medline: 29016974]
8. Ekblaw A, Azaria A, Halamka JD, Lippman A. Semantic Scholar. 2016. A Case Study for Blockchain in Healthcare: 'MedRec' Prototype for Electronic Health Records and Medical Research Data URL: https://www.semanticscholar.org/paper/A-Case-Study-for-Blockchain-in-Healthcare-%3A-%E2%80%9C-%E2%80%9D-for-Ekblaw-Azaria/56e65b469cad2f3ebd560b3a10e7346780f4ab0a
9. da Conceicao AF, da Silva FS, Rocha V, Locoro A, Barguil JM. arXiv. 2018. Eletronic Health Records using Blockchain Technology URL: https://arxiv.org/pdf/1804.10078

10. Roehrs A, da Costa CA, da Rosa RR. OmniPHR: a distributed architecture model to integrate personal health records. J Biomed Inform 2017 Dec;71:70-81 [FREE Full text] [doi: 10.1016/j.jbi.2017.05.012] [Medline: 28545835]
11. Ribitzky R, Clair J, Houlding DI, McFarlane CT, Ahier B, Gould M, et al. Pragmatic, interdisciplinary perspectives on blockchain and distributed ledger technology: paving the future for healthcare. Blockchain Healthc Today 2018 Mar 23;1:1-15. [doi: 10.30953/bhty.v1.24]
12. Zhang P, Schmidt DC, White J, Lenz G. Blockchain technology use cases in healthcare. In: Advances in Computers. First Edition. Cambridge, Massachusetts: Academic Press; 2018:1-41.
13. Hevner A, Chatterjee S. Design Research in Information Systems: Theory and Practice. New York: Springer; 2010.
14. Hevner AR. A three cycle view of design science research. Scand J Inf Syst 2007;19(2):87-92 [FREE Full text]
15. Hevner A, March ST, Park J, Ram S. Design science in information systems research. MIS Q 2004;28(1):75-105. [doi: 10.2307/25148625]
16. vom Brocke J, Simons A, Niehaves B, Riemer K, Plattfaut R, Cleven A. Reconstructing the Giant: On the Importance of Rigour in Documenting the Literature Search Process. In: Proceedings of the 17th European Conference on Information Systems. 2009 Presented at: ECIS'09; June 8-10, 2009; Verona, Italy p. 2206-2217.
17. George JF, Kohnke E. Personal health record systems as boundary objects. Commun Assoc Inf Syst 2018;42(1):21-50. [doi: 10.17705/1CAIS.04202]
18. Tasatanattakool P, Techapanupreeda C. Blockchain: Challenges and Applications. In: Proceedings of the International Conference on Information Networking. 2018 Presented at: ICOIN'18; January 10-12, 2018; Chiang Mai, Thailand p. 473-475. [doi: 10.1109/ICOIN.2018.8343163]
19. Myers MD, Newman M. The qualitative interview in IS research: examining the craft. Inf Organ 2007 Jan;17(1):2-26. [doi: 10.1016/j.infoandorg.2006.11.001]
20. Zhang P, White J, Schmidt DC, Lenz G, Rosenbloom ST. FHIRChain: applying blockchain to securely and scalably share clinical data. Comput Struct Biotechnol J 2018;16:267-278 [FREE Full text] [doi: 10.1016/j.csbj.2018.07.004] [Medline: 30108685]
21. Yang G, Li C. A Design of Blockchain-Based Architecture for the Security of Electronic Health Record (EHR) Systems. In: Proceedings of the International Conference on Cloud Computing Technology and Science. 2018 Presented at: CloudCom'18; December 10-13, 2018; Nicosia, Cyprus. [doi: 10.1109/CloudCom2018.2018.00058]
22. Gazzarata G, Gazzarata R, Giacomini M. A standardized SOA based solution to guarantee the secure access to EHR. Procedia Comput Sci 2015;64:1124-1129. [doi: 10.1016/j.procs.2015.08.582]
23. Badr S, Gomaa I, Abd-Elrahman E. Multi-tier blockchain framework for IOT-EHRS systems. Procedia Comput Sci 2018;141:159-166. [doi: 10.1016/j.procs.2018.10.162]
24. Mertz L. (Block) chain reaction: a blockchain revolution sweeps into health care, offering the possibility for a much-needed data solution. IEEE Pulse 2018;9(3):4-7. [doi: 10.1109/MPUL.2018.2814879] [Medline: 29757744]
25. Saitwal H, Feng X, Walji M, Patel V, Zhang J. Assessing performance of an electronic health record (EHR) using cognitive task analysis. Int J Med Inform 2010 Jul;79(7):501-506. [doi: 10.1016/j.ijmedinf.2010.04.001] [Medline: 20452274]
26. Künzi J, Koster P, Petkovic M. Emergency access to protected health records. In: Adlassnig KP, Blobel B, Masic I, Mantas J, editors. Medical Informatics in a United and Healthy Europe. Amsterdam: IOS Press; 2009:705-709.
27. Bossen C, Jensen LG, Udsen FW. Evaluation of a comprehensive EHR based on the DeLone and McLean model for IS success: approach, results, and success factors. Int J Med Inform 2013 Oct;82(10):940-953. [doi: 10.1016/j.ijmedinf.2013.05.010] [Medline: 23827768]
28. Fragidis LL, Chatzoglou PD. Implementation of a nationwide electronic health record (EHR): the international experience in 13 countries. Int J Health Care Qual Assur 2018 Mar 12;31(2):116-130. [doi: 10.1108/IJHCQA-09-2016-0136] [Medline: 29504871]
29. Peterson K, Deeduvanu R, Kanjamala P, Boles K. Office of the National Coordinator for Health Information Technology. 2016. A Blockchain-Based Approach to Health Information Exchange Networks URL: https://www.healthit.gov/sites/default/files/12-55-blockchain-based-approach-final.pdf
30. Haas P. Bertelsmann Stiftung. 2017. Elektronische Patientenakten URL: https://www.bertelsmann-stiftung.de/fileadmin/files/BSt/Publikationen/GrauePublikationen/VV_eEPA_Expertise_final.pdf
31. Weiskopf NG, Weng C. Methods and dimensions of electronic health record data quality assessment: enabling reuse for clinical research. J Am Med Inform Assoc 2013 Jan 1;20(1):144-151 [FREE Full text] [doi: 10.1136/amiajnl-2011-000681] [Medline: 22733976]
32. Marnau N. Die Blockchain im Spannungsfeld der Grundsätze der Datenschutzgrundverordnung. In: Proceedings of the 47th Annual Meeting of the Gesellschaft für Informatik. 2017 Presented at: INFORMATICS'17; September 25-29, 2017; Chemnitz, Germany. [doi: 10.18420/in2017_105]
33. Hopwood D, Bowe S, Hornby T, Wilcox N. This script - GitHub. 2019. Zcash Protocol Specification URL: https://raw.githubusercontent.com/zcash/zips/master/protocol/protocol.pdf
34. Ølnes S. Beyond Bitcoin Enabling Smart Government Using Blockchain Technology. In: Proceedings of the International Conference on Electronic Government. 2016 Presented at: EGOV'16; September 4-7, 2016; Petersburg, Russia p. 253-264. [doi: 10.1007/978-3-319-44421-5_20]

35. Hanseth O, Lyytinen K. Design theory for dynamic complexity in information infrastructures: the case of building internet. J Inf Technol 2010 Mar;25(1):1-19. [doi: 10.1057/jit.2009.19]
36. Scott B. EconStor. 2016. How Can Cryptocurrency and Blockchain Technology Play a Role in Building Social and Solidarity Finance? URL: https://www.econstor.eu/bitstream/10419/148750/1/861287290.pdf
37. Dagher GG, Mohler J, Milojkovic M, Marella PB. Ancile: privacy-preserving framework for access control and interoperability of electronic health records using blockchain technology. Sustain Cities Soc 2018 May;39:283-297. [doi: 10.1016/j.scs.2018.02.014]
38. Shabani M. Blockchain-based platforms for genomic data sharing: a de-centralized approach in response to the governance problems? J Am Med Inform Assoc 2019 Jan 1;26(1):76-80. [doi: 10.1093/jamia/ocy149] [Medline: 30496430]
39. Thwin TT, Vasupongayya S. Blockchain Based Secret-Data Sharing Model for Personal Health Record System. In: Proceedings of the 5th International Conference on Advanced Informatics: Concept Theory and Applications. 2018 Presented at: ICAICTA'18; August 14-17, 2018; Krabi, Thailand p. 196-201. [doi: 10.1109/ICAICTA.2018.8541296]
40. Croman K, Decker C, Eyal I, Gencer AE, Juels A, Kosba A, et al. On Scaling Decentralized Blockchains. In: Proceedings of the International Conference on Financial Cryptography and Data Security. 2016 Presented at: FC'16; February 22-26, 2016; Christ Church, Barbados p. 106-125. [doi: 10.1007/978-3-662-53357-4_8]
41. Kosba A, Miller A, Shi E, Wen Z, Papamanthou C. Hawk: The Blockchain Model of Cryptography and Privacy-Preserving Smart Contracts. In: Proceedings of the Symposium on Security and Privacy. 2016 Presented at: SP'16; May 22-26, 2016; San Jose, CA, USA p. 839-858. [doi: 10.1109/SP.2016.55]

Abbreviations

DSR: design science research
EHR: electronic health record
ICTs: information and communication technologies
ISO: International Standard Organization
IT: information technology
KB: key benefit
KC: key challenge
RQ: research question

Edited by K Clauson, P Zhang; submitted 01.02.19; peer-reviewed by TT Kuo, M Bestek, S Sarbadhikari; comments to author 31.03.19; accepted 27.04.19; published 07.10.19.

Please cite as:
Beinke JH, Fitte C, Teuteberg F
Towards a Stakeholder-Oriented Blockchain-Based Architecture for Electronic Health Records: Design Science Research Study
J Med Internet Res 2019;21(10):e13585
URL: https://www.jmir.org/2019/10/e13585
doi:10.2196/13585
PMID:31593548

Tutorial

Building a Secure Biomedical Data Sharing Decentralized App (DApp): Tutorial

Matthew Johnson[1,2], MS; Michael Jones[1,2], MS; Mark Shervey[1,2], BA; Joel T Dudley[1,2], PhD; Noah Zimmerman[1,2], PhD

[1]Center for Biomedical Blockchain Research, Icahn School of Medicine at Mount Sinai, Redwood City, CA, United States

[2]Institute for Next Generation Healthcare, Icahn School of Medicine at Mount Sinai, New York City, NY, United States

Corresponding Author:
Noah Zimmerman, PhD
Center for Biomedical Blockchain Research
Icahn School of Medicine at Mount Sinai
Suite 10
234 Marshall St
Redwood City, CA
United States
Phone: 1 650 352 3879
Email: noah.zimmerman@mssm.edu

Abstract

Decentralized apps (DApps) are computer programs that run on a distributed computing system, such as a blockchain network. Unlike the client-server architecture that powers most internet apps, DApps that are integrated with a blockchain network can execute app logic that is guaranteed to be transparent, verifiable, and immutable. This new paradigm has a number of unique properties that are attractive to the biomedical and health care communities. However, instructional resources are scarcely available for biomedical software developers to begin building DApps on a blockchain. Such apps require new ways of thinking about how to build, maintain, and deploy software. This tutorial serves as a complete working prototype of a DApp, motivated by a real use case in biomedical research requiring data privacy. We describe the architecture of a DApp, the implementation details of a smart contract, a sample iPhone operating system (iOS) DApp that interacts with the smart contract, and the development tools and libraries necessary to get started. The code necessary to recreate the app is publicly available.

(J Med Internet Res 2019;21(10):e13601) doi:10.2196/13601

KEYWORDS

blockchain; geolocation; tutorial; mobile health; privacy; DApp; iOS; biomedical research; decentralized application; smart contract

Introduction

Background

Decentralized apps (DApps) are computer programs that run on a distributed computing system. These have been popularized recently by distributed ledger technologies underlying projects such as Bitcoin and Ethereum. Unlike the client-server architecture that powers most internet apps, DApps interact with app logic deployed on a blockchain enabling transparent, verifiable, and immutable records of each transaction. When built on blockchain networks, DApps can contain their own suite of associated smart contracts that are used to encode business logic and allow persistent storage of state [1].

Over 150 blockchain projects in the health care industry alone have raised more than US $660 million in private and blockchain-funded markets [2,3]. Despite this massive investment in blockchain technologies over the last 3 years, the DApp ecosystem remains immature. At the time of writing, the most popular DApp had 5628 daily active users [4]. By comparison, Facebook, a popular centralized app, reported 1.5 billion daily active users in Q4 of 2018 [5]. Documentation and tooling for developers to build on new platforms is sparse [6,7]. As a result, the technical hurdles currently required to develop a DApp restrict more widespread experimentation.

Self-contained sample projects [6-10] can help jump-start development, allowing those new to the field to focus on the logic and app instead of the myriad technical decisions required to get started. This tutorial too walks through the development

of a complete working DApp prototype, including smart contract and iPhone operating system (iOS) client apps, and is specifically motivated by a real use case that requires data privacy in biomedical research. All of the code necessary to recreate this app has been made publicly available [11,12]. Our hope is that this project will be forked, remixed, and combined to catalyze the ecosystem of biomedical blockchain DApp development.

This tutorial is organized as follows: (1) a discussion around the traditional approach to data privacy in research as well as the rationale for a blockchain approach; (2) the motivation for a geolocation sharing use case for biomedical research; (3) an overview of the software architecture proposed, including a brief description of security properties; (4) a description of the client app; (5) a description of the smart contract; (6) details on the development environment; (7) details on the deployment of the DApp; and (8) a discussion of advantages, limitations, and future directions.

The aim of this paper is to serve as a tutorial for developing a working prototype of a DApp (as shown in the demo video in Multimedia Appendix 1) and to highlight the specific benefits of using a DApp as a method for participants in research to share features of their raw data, while preserving the participant's privacy by not revealing the raw data itself.

Traditional Data Protection and Its Shortcomings

Traditionally, data collected in research is managed using a database that is implemented in a client-server network architecture. In this architecture, a database is connected to a backend server, which can then be accessed by researchers (clients). An example of this architecture is an iOS app, functioning as the front end (client), which makes calls to backend code and the database (server) [13]. This database requires some central authority, typically the researcher or research organization, to set permissions and control access resulting in a centralized app. It is the responsibility of the researchers to ensure the privacy and protection of the collected data; this architecture is convenient for researchers in that regard, as they technically have full autonomy over all the data they collect.

Researchers having full access and control over all collected data from participants, however, is not necessarily in the best interest of either party. This approach can place undue burden on researchers who are only interested in nonidentifiable features of the data and would rather not bear the liability of collecting and managing data that is unnecessary to their analysis. On the contrary, participants must blindly trust that the researchers will responsibly manage and protect their raw data, which, by malice or negligence, is not always the case. This becomes increasingly concerning as longitudinal research and high-frequency data collection is becoming more prevalent, exposing participants to greater privacy risks [14].

Although the reduction of data collection and management liability is realizable in any traditional hosted environment, the party hosting the environment will still have access to the raw data. To avoid this, one could employ data minimization and perform all data-processing and feature extraction on the client so that the host never has access to the raw data itself.

For example, researchers could be interested in the points in time that the heart rate of participants went above 60 beats per minute (bpm). Rather than recording, posting, and storing all the heart rate data of participants, the heart rate could be preprocessed on the participant's mobile device for heart rate data above 60 bpm. However, this would eliminate the ability for researchers or future collaborators who are interested in heart rate data above 70 bpm to use the previously collected data as there is no way to determine the heart rates for the previously collected time points. To collect new heart rate data above 70 bpm would require an update to the client app.

Unlike the traditional, centralized approach of data privacy where the backend code and database are on a centralized server, DApps have backend code running on a decentralized peer-to-peer network, such as a blockchain network [13]. When deployed on a blockchain, the backend code is referred to as a smart contract. A smart contract is what the client uses to interact with a blockchain network. DApps that are built on smart contracts offer unique advantages over centralized client-server alternatives [15] and in the right situations can serve as a trusted intermediary for data management between participants and researchers. As further detailed in Table 1, transparency, autonomy, immutability, and self-sufficiency are some of the core features of blockchain technology and therefore DApps.

As smart contracts are transparent and immutable, it is possible to verify and guarantee the behavior when the smart contract source code is made publicly available by comparing the binary code of the deployed contract against the open-source code. This may, in part, alleviate the participants' concern that their data could be mishandled or mismanaged. On the contrary, there are no guarantees that traditional database implementations do not secretly collect and log information, or that the data access policies remain private for all time.

Researchers can rely on smart contracts to self-execute for example, to return relevant features of interest from participant data without the researcher having any awareness of the raw data. This allows for the collection of more raw data, while still limiting the exposure of these data to the researcher. In the heart rate example mentioned earlier, all of the heart rate data could be posted to a smart contract, which can be written in such a way that would allow participants to control access to these data. For example, participants could update researcher access from limiting access to heart rate data above 70 bpm to allowing access for heart rate data above 60 bpm. Using only data minimization, researchers would not be able to access previous heart rate data above 60 bpm. Using smart contracts, researchers could achieve data minimization and reduction of liability, while maintaining a degree of flexibility when accessing private data.

Table 1. Smart contract properties: benefits and trade-offs.

Property	Description	Benefits	Trade-offs
Transparent	The state of the app is public and inspectable.	App functionality can be audited and validated; Public nature of code incites collaboration.	Requires careful implementation to avoid exposure of sensitive data; Vulnerabilities can more easily be identified and exploited.
Autonomous	Can be designed and deployed such that it does not require any further interaction with the agent that deployed it.	No need for middleman or external arbiter. No external control or manipulation of app behavior.	No customer service: transactions cannot be reversed, and corrections cannot be made.
Immutable	The code defining the contract cannot be modified.	Guarantees that data policy will not change.	Cannot update smart contract with security fixes; requires new contract deployment.
Self-sufficient	Has the ability to coordinate and incentivize resources, in the form of tokens, to execute functions.	A deployed contract stays deployed on blockchain; does not require developer to pay for maintaining a server.	Lack of control over deployed malicious contracts; users of the smart contract pay transaction costs.

It is important to note that there are several alternative approaches that strive to provide guarantees of data privacy preservation. Many of these approaches rely on trusted third parties, cryptographic, secure hardware, or blockchain-based techniques and are actively being researched and are under development. A full examination of these approaches is out of scope for this tutorial, but a forthcoming article that accompanies this study provides a detailed comparison of techniques based on ability to preserve data privacy and on the practicality of implementation [16]. One of the primary findings is that blockchain-based techniques, particularly when combined with other techniques such as secure hardware, can provide high levels of data privacy, verifiability, and practicality of implementation.

Use Case: Sharing Location Data for Biomedical Research

This tutorial presents a mobile iOS DApp that allows a research study participant (*participant*) to share useful features of their location data derived from global positioning system (GPS) coordinates, or geocoordinates, with a research team (*third party*), without revealing their raw GPS coordinates. Location data have proven diversely useful in biomedical research where it has improved disease management and treatment delivery, been used to monitor behavioral and environmental risk factors, and even has informed public health policy in substance abuse [17-20].

Although geolocation data hold significant promise for a variety of health care apps, location data are also one of the most sensitive pieces of personal information [21]. It is in the interest of both the researchers and the research participants that the collection of these data has been restricted to the stated goal of the research project, and this study aims to show that blockchain technology can be deployed in such a way to attain these ends.

Decentralized App Architecture and Use

The DApp consists of a *client app* for the participant and the third-party researcher and a *smart contract* (Figure 1). A smart contract is a collection of code and data that encapsulate the business logic of the DApp [1]. Smart contracts are written in a high-level programming language, compiled into bytecode, and deployed to a unique address on a blockchain. The development and execution of smart contracts are supported by a variety of blockchain platforms such as Ethereum, EOS, Tron, POA, and Oasis [4,22]. Each blockchain implementation offers different features and trade-offs based on their purpose and protocol; therefore, it is important to consider the features when considering a blockchain platform solution. For additional blockchain related terms and definitions mentioned throughout this study, see Multimedia Appendix 2.

As shown in Table 2, we compare features of a traditional database against Ethereum, a popular public open-source blockchain platform, and against Oasis Devnet, a privacy-preserving blockchain solution.

Data privacy is critical when dealing with biomedical data. However, most public blockchains lack confidentiality and privacy of state variables and data. Smart contracts deployed on Ethereum, for example, allow for the state and data stored within them to be read by anyone. Therefore, public blockchains that lack data privacy and confidentiality would not be a suitable standalone solution for storing and handling sensitive biomedical data [23,24].

The Oasis Devnet addresses this critical gap in blockchain's lack of confidentiality by combining blockchains with trusted execution environments (TEEs) [25]. The Oasis Devnet is based on Ekiden, a system anchored in a formal security model expressed as a cryptographic ideal functionality [25,26]. The underlying blockchain in Oasis is encrypted, which prevents the dedicated storage of the contract data and state to be read, unlike public blockchains. Within Ekiden, anyone with a TEE-enabled platform can participate as a compute node to execute contracts within the contract TEE. To create or retrieve the keys required to run the contract, the contract TEE must reach out to the key management committee, a quorum of compute nodes that manage the keys needed to run the contract. This system prevents a malicious node from forking the blockchain and acting as a compute node, as they would not have the necessary keys from the key management committee to run the contract.

Additional technical details involving the security and privacy-preserving features underlying the Oasis platform can be found in the study by Cheng et al [25].

Therefore, we selected the Oasis Devnet to deploy this DApp because it (1) prioritizes privacy preservation, (2) has a functional and supported developer network, and (3) is compatible with the Ethereum toolkit.

Figure 1. Decentralized application architecture and workflow—Smart contracts consist of self-executing code run on a blockchain protocol. Data flow directly between the smart contract and the clients: (1) Participant submits timestamped geolocation data; (2) Participant grants/revokes permission to share that data, to the smart contract; (3) A third party assigns geolocations of interest a matching category (ie, hospital, gym, pharmacy, or none) and deploy that mapping to the smart contract; (4) Participant can view the timestamp of each of their previously written geolocations and the category of that geolocation, if there exists a mapping between that geolocation and a category that was previously written to the smart contract by a third party; and (5) A third party can view timestamped data that the participant has chosen to share.

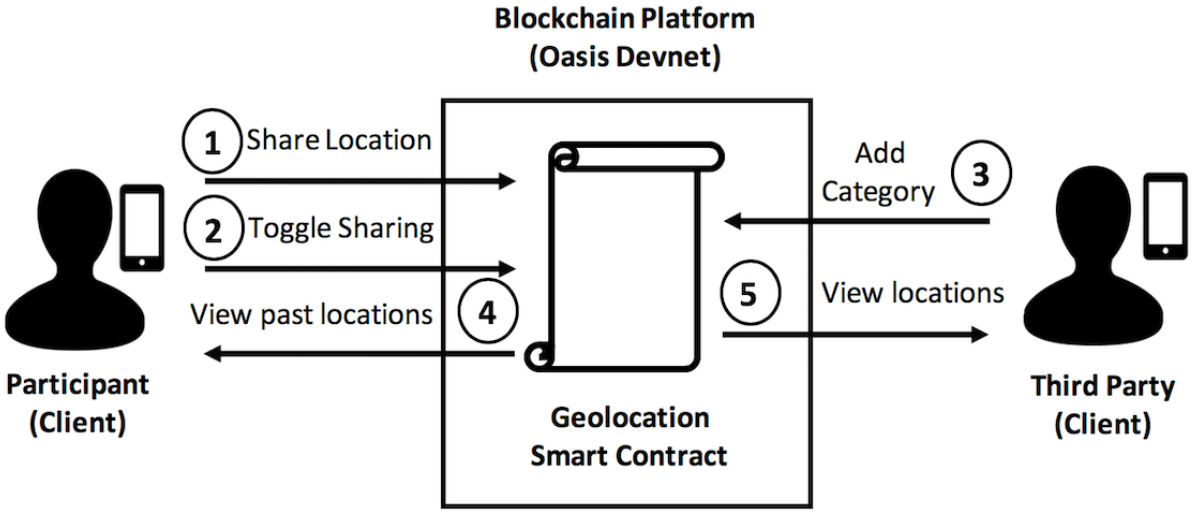


Table 2. Features of traditional databases compared with Ethereum and the Oasis Devnet.

Features	Traditional database	Ethereum (public)	Oasis Devnet
Data read access control	Yes	No, data is public	Yes, dependent on smart contract logic
Anonymity	Yes, if host is honest	Pseudo-anonymous	Pseudo-anonymous
Cost	Fixed	Variable	Free on developer network, will be variable in production mainnet
Data privacy	Yes, if host is honest	No, data is public	Yes, dependent on smart contract logic
Data mutability	Mutable, but can be immutable via role permissions	Immutable	Devnet gets reset, but the production mainnet will be immutable
Code can be updated	Yes	Yes	Not yet. Intercontract calls are planned
Publicly verifiable	No, the public cannot verify stored procedures	Yes, the public can verify smart contract code[a]	Yes, the public can verify smart contract code[a]
Widely accessible	Yes	Yes	Yes

[a]For both Ethereum and Oasis Devnet, the smart contract source code must be made public to verify that the contract is doing what is claimed.

The Oasis Devnet supports several runtimes, allowing developers to choose between popular languages. Here we use Solidity, an object-oriented programming language for writing and implementing smart contracts, because of its existing development frameworks and interoperability with Ethereum, the second largest blockchain network [27]. It should be noted that the Oasis Devnet was used, as the mainnet was not yet available at the time of this writing. In contrast to an open testnet, the Oasis Devnet is hosted by Oasis Labs and was created specifically for developers to make an easy-to-use, developer friendly environment to develop and test on Oasis [28].

Using a blockchain that supports private states, such as Oasis, allows the handling of raw data to shift from that of a central party to a blockchain platform that provides a combination of transparency and privacy via verifiable smart contracts and a system by which no party can view or access the raw data.

Client App

To simplify the tutorial codebase, the iOS app can toggle between a participant mode and a third-party mode; in practice the two views would be implemented as separate apps that only display the functionality relevant to that user type. Multiple participant users and third-party researcher users could use the client app at once, depending on the desired enrollment and collaboration goals of the research study design.

Participant Mode

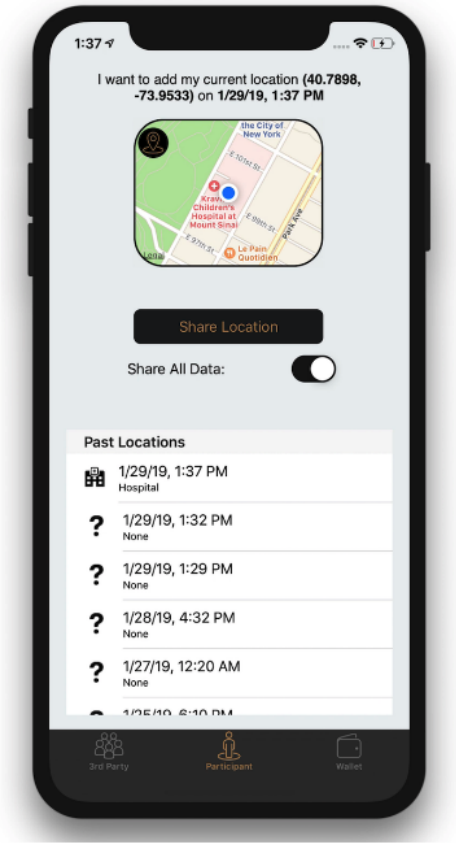
The participant tab presents a participant user with the common functionalities and user interface that a participant in a research study would expect (Figure 2). These include posting the user's current location, toggling data sharing, and viewing the user's previously posted data.

This view also presents the user with the device's geocoordinates and local time to display the human-readable data that are posted to the smart contract, should the user decide to share their location. The geocoordinates are represented as the latitude and longitude of the device with a precision of 4 decimal degrees, approximately 11.132 meters at the equator. This precision was chosen to provide sufficient granularity between establishments, without requiring several adjacent coordinates to represent a single establishment. When posting to the smart contract, these geocoordinates are multiplied by a factor of 10^4 to represent the decimal degrees as signed integers. If the user has not previously posted a location, a new participant identifier (ID) for the user is created, and the associated sharing status is enabled and set to true.

The participant can toggle whether or not to continue sharing their current and previous data anytime. If sharing is enabled then all third parties are able to view the number of previously submitted locations by the participant, along with the timestamped category. If sharing is disabled, third parties can only see that participant's ID and that they have chosen not to share data at this time. A design decision was made to make the default sharing status opt-out, for 3 reasons: (1) The primary participant user action is to "share location," which itself informs the participant that their data will be shared; (2) Participant data privacy is provided regardless of the sharing status, but is available because we believe participants own their data and should have the ability to revoke access; and (3) To maximize benefit to third parties who want to access to participant location feature data.

A third party can benefit from seeing the ID of participants that have revoked data access to better understand how many participants exist and how many of them continue to share data over time. Showing all participant IDs was included primarily as a demonstration to easily show how access can be granted and revoked; however, the smart contract could be easily written in such a way that no longer allows third parties visibility into the participant IDs who have revoked access. This may be an important security consideration in scenarios where there are few participants. As additional participant IDs are revealed to third parties, regardless of sharing status, this could provide some level of data to malicious third parties that may attempt to extrapolate a participant's wallet address from the participant ID by examining the history of transactions with the smart contract. Fortunately, the method that is called by a transaction is concealed, so there is no direct means of differentiating between a participant and a third party based on examining the transaction history.

Figure 2. Simulator running the iPhone operating system app displaying the participant mode.



The participant is also presented with a table containing previously posted locations ordered by the time posted. Each location displays its corresponding category and timestamp. These previously posted locations will always be visible to the participant that posted them, even if sharing is disabled at that time. This allows a user to view the data that will be shared if they choose to do so.

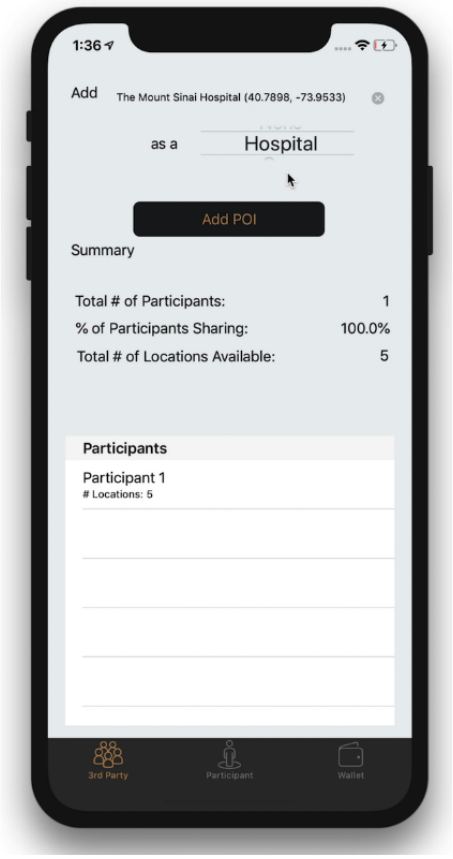
Third-Party Mode

The third-party view presents views and functionality for a third party, such as a research coordinator (Figure 3). These include categorizing a geolocation and viewing participant feature data.

Within this view, a third-party user is able to search for map-based addresses and places of interest. The places of interest are centered around the location of the device, through Apple's MapKit Framework, where the geoencoding service is performed by Apple. It should be noted that the geoencoding service is done solely as a convenience to the third party and is not done for the participant nor identifies the participant's location. Once the user has selected a location and selected its associated category, they can post to the smart contract.

The view also presents a summary of participant data, such as the number of participants who have posted at least one location, the percentage of participants who have enabled sharing, as well as the total number of locations that are available and being shared. For specific information on each participant, the third party is shown a table view of all participants and whether the participant has enabled the sharing of their data. If a participant has enabled the sharing of their data, third parties are able to view how many locations a participant has posted and can view the timestamp and location category of each entry.

Figure 3. Simulator running the iPhone operating system app displaying the third-party mode.



Smart Contract

The smart contract manages participant enrollment in the research study and allows participants to share their geocoordinates with the study. The smart contract also determines whether the geocoordinates correspond to a predefined location type and whether or not to allow third parties to view the data participants commit to the smart contact. Importantly, raw geolocation data are not stored within the smart contract, nor are the geolocation data directly linked to the participant.

Decentralized App Development

Tools and Libraries

Many tools and libraries exist for the development of DApps. The tools and libraries used to develop this DApp are briefly introduced and summarized and will be expanded upon further throughout the discussion of development (Table 3). Certain resources in this tutorial are platform-dependent, such as Oasis Contract Kit, whereas others are generally useful for DApp development such as ConsenSys AG's MetaMask and Truffle.

Table 3. Decentralized app–related resources used in this tutorial.

Tools and Libraries	Descriptions
MetaMask	Browser extension that serves as a Web3 wallet that can create and manage identities. It also injects the web3.js library into the browser to allow read and write requests to be made on blockchain networks, such as Ethereum or other networks by specifying a remote procedure call URL.
Oasis Contract Kit	Docker environment with a preconfigured set of tools for developing contracts on Oasis.
Oasis Devnet	Privacy-focused blockchain platform for developers to build and test confidential smart contracts; the platform used in this tutorial.
Remix	Web browser–based integrated development environment that allows developers to write, deploy, and run smart contracts written in Solidity.
Solidity	An object-oriented programming language for writing and implementing smart contracts on various blockchain platforms.
Truffle Framework	A development environment testing framework and asset pipeline for blockchains using the Ethereum Virtual Machine; included in Contract Kit.
Web3swift	Open-source iOS library written in Swift. It provides web3.js functionality in Swift, native ABI parsing, and smart contract interactions.

Client App

The client is a native iOS app, written in Swift 4.2 using XCode 10.1. The iOS app interacts with the smart contract deployed on the Oasis Devnet via the open-source web3swift library written by Matter, Inc. This library allows for the interaction with a remote node of the Oasis Devnet via JavaScript object notation (JSON) remote procedure call (RPC) as well as smart contract interaction [29]. The web3swift library also provides local keystore management, which assists the app user in creating and importing a wallet as well as the creation and management of public/private keys. This eliminates an additional burden and barrier to entry for users unfamiliar with wallets and key management. When implementing the web3swift library, the web3 instance was bound to the RPC URL provided in the Oasis documentation [22]. The contract instance within the iOS app was initialized from the app binary interface, or ABI, string in JSON format and was obtained from the Ethereum Foundation's Remix integrated development environment (IDE) or whichever IDE was used to develop the smart contract. The ABI is a data encoding and decoding scheme and is the standard way to interact with smart contracts in Ethereum for interfacing with smart contracts.

It should be noted that as the client app is a proof-of-concept, there are many improvements which can and should be made in production. Currently, none of the data read from the smart contract are persistently stored on the device, which hinders performance. Similarly, none of the raw geolocation data are persistently stored on the device. If any data were to be persistently stored, it should be encrypted. Currently, the only data that are persistently stored on device includes the public wallet address, whether that address has been registered, and the encrypted private key of said address with a user provided password via AES-128-CBC. In practice, access to data should require authentication whenever the app is left in an inactive state for a given amount of time.

Smart Contract

The smart contract is developed in Remix, a Web browser–based IDE that allows developers to write, deploy, and run smart contracts written in Solidity [30]. Remix allows for contract deployment in various types of environments, including: a JavaScript virtual machine (VM) in which transactions are executed in a sandbox blockchain in the browser, an Injected Provider in which Remix will connect to an injected Web3 provider such as MetaMask, and finally a Web3 provider in which Remix will connect to a remote blockchain node [31]. Initially, the smart contract was developed in the JavaScript VM environment in Remix for its simplicity.

Figure 4 illustrates how raw geolocation data are not stored within the smart contract and also not directly linked to the participant. This design was selected because Oasis Devnet offers Ethereum backward compatibility with support for existing Ethereum wallets. This allows for existing Ethereum wallets to be imported into the app or for newly created wallets within the app to also be used on Ethereum. Although convenient, this could result in the exposure of identifying information on the Ethereum network being effortlessly correlated to the same address on the Oasis Devnet, should the same address be used on both networks. Therefore, a participant's wallet address is not revealed to third parties, and instead, third parties have access to the participant ID.

Figure 4. Solidity method written to post the location of the participant. This method (postParticipantLocation), when called by the client, (1) checks to see if the sender is a new participant by checking to see if there exists a mapping for the sender's address to an existing participant ID (addressToParticipant). (2) If one does not exist, the total number of participants is incremented (numberOfParticipants). (3) The senders address is then mapped to a participant ID that is equal to the count of participants.

```
/// @notice Posts location of msg.sender as a participant
/// @param _dateTime Unix time participant was at location
/// @param _lat Latitude of the participant rounded to the nearest 4 decimal degrees and multiplied by 10^4
/// @param _long Longitude of the participant rounded to the nearest 4 decimal degrees and multiplied by 10^4
function postParticipantLocation(uint256 _dateTime, int256 _lat, int256 _long) public {

    if (addressToParticipantID[msg.sender] == 0) {                  // Check if msg.sender is a new Participant
        numberOfParticipants ++;                                    // Increment total number of participants
        addressToParticipantID[msg.sender] = numberOfParticipants;  // Assign the new participant a new ParticipantID
        sharingEnabled[numberOfParticipants] = true;                // Enable sharing of data
    }

    uint256 participantID = addressToParticipantID[msg.sender];     // Get ParticipantID
    bytes32 curLoc = keccak256(abi.encodePacked(_lat, _long));      // Create a hash of the latitude and longitude
    participantCoordinates[participantID][_dateTime] = curLoc;      // Map geolocation hash to dateTime of ParticpantID
    participantDateTimes[participantID].push(_dateTime);            // Append new dateTime to PaticipantID
}
```

Considering that geolocation data should not be directly linked to the participant, the participant ID is used to identify a participant. Once the sender address has been assigned and mapped to a participant ID, the keccak256 function is used to compute the Ethereum-SHA-3 (Keccak-256) hash of the latitude and longitude input parameters (*_lat, _long*) as a convenient encoding. This hash is stored as a private variable, not accessible outside of the smart contract, and later used to uniquely identify this particular geolocation within the smart contract. The hashed geolocation is mapped to the timestamp input parameter (*_dateTime*), which is subsequently mapped to the respective participant ID. This hashed geolocation is also used to compare with other hashed geolocations, to determine if that particular geolocation matches a category type posted by a third party. The timestamp input parameter (*_dateTime*) is then appended to a mapped array of the participant ID (*participantDateTimes*). This allows third parties to assign a category to a particular geolocation. It is important to note that the privacy of these input variables, hashed geolocations, and mappings within the smart contract are made possible by the Oasis Devnet as it supports private data that can only be accessed by the smart contract themselves. Descriptions of additional variables and functions within the smart contract can be found in Multimedia Appendices 3 and 4.

A similar method exists for both storing a mapping of the particular category type to the Ethereum-SHA-3 (Keccak-256) hash of the geolocations and storing a mapping of an array of the hashed locations to the category type.

It is worth noting that these category types were deliberately predefined as *enum* types on contract creation to prevent third parties from creating and storing new and custom categories that could be used to identify particular locations (eg, by posting a custom category that labels the accompanying geolocation with a postal address or geocoordinate pair.) This provides a way for third parties to add new locations that match the predefined category types into the future, whereas at the same time holding them responsible for the transaction cost of adding these new locations.

Decentralized App Deployment

Wallet and Funding

Before the deployment of the smart contract onto the Oasis Devnet, a hierarchical deterministic (HD) wallet was created and funded from the Oasis faucet. MetaMask was installed and set up to be used as this wallet [32]. When creating the wallet, the mnemonic phrase, or seed words, used to generate the HD wallet was safely stored and made easily accessible, as it was later needed for deploying the contract. Once the wallet and account were created, the network within MetaMask was changed to the custom RPC URL, chainID, symbol, and nickname provided by Oasis Labs to connect MetaMask to the Oasis Devnet [28].

Funds were acquired via the Oasis Devnet Faucet by following the onscreen instructions [33]. Once the funds were received, the wallet within MetaMask updated to show the amount funded in DEV. DEV, the symbol for Oasis Devnet tokens, were used to pay the transaction fees needed to deploy the contract [22]. DEV have no value and work only on the Oasis Devnet, as they will not transfer to the Oasis mainnet [22]. It was important that the first account created by MetaMask was the one that was funded, otherwise deployment of the contract to the Oasis Devnet would have failed. If a different account was funded by mistake, DEV could have been transferred from the funded account to the first account.

Contract Deployment

The contract was deployed using Oasis Contract Kit. The Oasis Contract Kit is a Docker environment with a preconfigured set of tools to provide developers with an environment that provides tools for developing, testing, and deploying confidential smart contracts. Following the steps provided by Oasis, the geolocation smart contract was deployed as both a confidential and nonconfidential smart contract on the Oasis Devnet (procedure described in Multimedia Appendix 5) [12,34].

iPhone Operating System App and Deployed Smart Contract

To interact with the nonconfidential smart contract, the iOS app was set, by default, to initialize a web3 instance using the RPC

URL provided by Oasis, as well as initialize an instance of the deployed smart contract using the ABI and contract address. Unlike Ethereum and other standard Web3 platforms, transactions to confidential smart contracts on the Oasis platform are encrypted end-to-end, where only the caller and the smart contract can decrypt transaction contents [35]. Communication with the Oasis Devnet occurs via hosted nodes by Oasis, and the client iOS app uses HTTPS, an encrypted communications protocol using Transport Layer Security, to protect data in transit. Oasis also provides a web3c.js client library that wraps underlying RPC calls with encryption and decryption, which allows one to securely communicate with smart contracts. Due to a lack of web3swift support for web3c, a web3 extension for confidential transactions, the iOS app was unable to interact with the deployed confidential smart contract at the time of this writing. Although the nonconfidential implementation is sufficient for the purposes of a working demonstration, it should be stressed that the nonconfidential implementation does not provide the participant any privacy of their posted location data.

Creating/Importing a Wallet

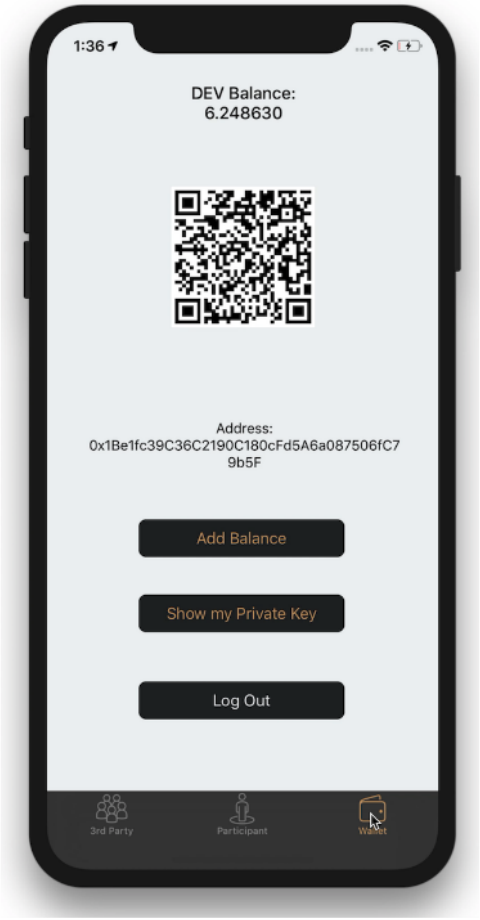
If there is no stored wallet on the first launch of the app, the user is presented with an option to create a new wallet or to import an existing wallet. When creating a wallet, the user is prompted to enter a user-generated password, which is later used to access the wallet private key and perform transactions that cost DEV, such as posting participant location data. Once set, the user-generated password cannot be recovered so it is recommended that it be written down or stored in a safe place. In the case of importing a wallet, the user would use their previously recorded password and would also have to provide their private key either manually keying it in or, for convenience, using the device's camera to scan a QR code encoded with the private key. Once a wallet is in place, the user will be presented with 3 tabs: *third party*, *participant*, and *wallet* [36].

Funding the Wallet

The wallet view presents the user with views and functionalities associated with managing a wallet on a blockchain and is required for both user types (Figure 5). These include allowing the user to view their wallet address and providing a link to fund the wallet. Regardless of whether the user is a third party or participant, the user would first have to add funds to their wallet, which is done within this tab. To make any changes to the state of the smart contract, such as a participant posting their geolocation or a third party assigning a new geolocation a location type, the user is still required to have funds to pay the transaction fees.

To add funds, a user would first navigate to the wallet tab to obtain their wallet address. For convenience, a button has been added that will copy the wallet address to the clipboard and open the URL for the Oasis Devnet Faucet in the default Web browser [33]. Once the user fills out their information to request funds for the Oasis Devnet, they will receive a confirmation email. Upon email confirmation, the user can return to view the updated balance of DEV for their wallet address. Within this tab, the user can also find a QR code with their embedded wallet address. As a security measure, any function resulting in a state change to the contract requires password reauthentication to gain access and view their wallet's private key. If the device is lost or the app deleted from the device, the only way to recover the wallet and access to the account is to import the private key.

Figure 5. Simulator running the iPhone operating system app, displaying the user's wallet information.



Discussion

To assess DApps and blockchain as a potential solution to the pitfalls of traditional data privacy preservation in research, a geolocation sharing use case was proposed. An iOS DApp and smart contract were developed as a proof-of-concept to illustrate the advantages and disadvantages that accompany this approach.

Although data privacy in research has typically been managed by a trusted third party, there are key areas in which this centralized approach falls short. Participants must trust third parties/researchers to properly use their data but also trust that their data are protected and secure from others who may misuse the data. As blockchains present an alternative method for data management and app logic that does not rely on a trusted third party, there is an opportunity where DApps could be explored as an alternative method for preserving data privacy.

Advantages

Smart Contract Properties

Smart contracts allow for transparency as the code can be made public and easily verified against the deployed smart contract (Table 1). As a result, users can trust that a validated smart contract does what it claims to do, such as not revealing raw location data in our case, proving it is able to act as a trusted intermediary for data management between participants and third parties. This public and open nature invites collaboration, allowing others to test the smart contract as well as share code of their own to improve the DApp community. Smart contracts are immutable; thus, participants can be confident that once a smart contract used by the DApp is verified and proven to be secure, the resulting management of data will also remain unchanged for the lifetime of the contract.

Costs

With traditional client-server architecture, there is a fixed cost in maintaining a backend database. Even if the database is no longer in use, or hardly ever accessed, there is still typically a cost for keeping it functioning and accessible. With a DApp, however, the only costs associated with hosting are the transaction cost of deploying the smart contract and the individual user costs incurred when modifying states within the contract.

An example of a state change in the example DApp is when a participant posts their location data to the smart contract. This is incurred because the smart contract is storing new data, and the state of the contract has changed. However, viewing data that exists in the smart contract can be done at no cost to a third party; for example, viewing how many participants are in a study. Depending on how the smart contract is written, viewing data could also come at a cost. This can be particularly useful in aligning and incentivizing different users toward a mutually beneficial goal. In the example DApp, the smart contract could be written in such a way that third parties pay a small fee, in DEV, to the participant, to view the category of each location

that the participant has visited. This may incentivize more participants to share more data with third parties.

Access Control

This DApp illustrates how a third party that is interested in a particular feature, such as a location category, could determine the feature without accessing any raw data. Limiting the scope of data shared with a third-party research team helps protect the participant's data from being misused outside the context of the study [14].

If, at any time and for any reason, the participant decides to stop sharing feature data with the third party, they are able to revoke third-party access to their data. As the source code governing the smart contract can be made public and can be verified to prove that the contract executes as claimed, participants can be assured that access to third parties is in fact revoked. Traditional databases, however, could continue to provide access to third parties unbeknown to a participant as access control cannot be publicly verified. Empowering the participant with control over their data may encourage participation and sharing of data by participants by explicitly addressing data privacy concerns.

Participant Identity

One advantage to using this DApp is the ability for the smart contract to create and maintain a private mapping between the participant ID and the participant's wallet address. This is advantageous as only the smart contract has access to this identifying and sensitive data, thus further safeguarding participants' identities. This also obviates the need for a separate party to manage the pseudonymization of the participants' identities.

Data Exposure

Traditional approaches may use external parties and services to perform feature extraction in an attempt to limit exposure to the raw location data. In this example, a third party may use a service, such as Google Places, to accomplish categorizing past locations. If such a service were used, that service could collect, store, and share a participant's raw geocoordinates and potentially identify the participant unbeknown to the third party or participant. Even if the raw location data is anonymized before using these external services, this increases privacy risks as it has been shown that deanonymization based on location data is possible [37,38]. This DApp provides an alternative to external services by potentially sourcing multiple interested third parties who are able to contribute to the mapping of locations to particular categories. Utilizing a DApp to perform feature extraction on the raw data provides one less potential entity in possession of participants' raw data.

Disadvantages and Limitations

Smart Contract Properties

Although there are several advantages to this DApp framework, there are also several limitations. Making smart contracts public and verifiable inherently allows for the discovery and exploitation of vulnerabilities. Should vulnerabilities be discovered, the smart contract cannot be updated in the same way a backend server would; instead it would require the deployment of a new smart contract. Smart contracts cannot be changed once deployed, so mistakes made by either a participant or third party cannot be reversed or corrected, except in the case where it was written into the smart contract. In blockchains that offer cross-contract calls, such as Ethereum, an upgradable smart contract design could be implemented. This design would offer the ability to deploy new smart contract versions by redirecting the contract address, thus allowing an upgradable contract. However, smart contracts within Oasis are unable to interact with one another at the time of this writing, not making it possible to implement in the context of the Oasis Devnet. In the future, however, Oasis plans to add support for cross-contract calls [39]. Additional design considerations during smart contract development can be found in Multimedia Appendix 6.

Transaction Costs

To make any state changes to the smart contract such as a participant posting their location or a third party posting a new location, the DApp user must pay a transaction fee. To pay this fee, the DApp user must have a minimum amount of funds in their wallet, which requires an additional task of obtaining funds. In the case of this DApp, the network is the Oasis Devnet, the fee is paid for in DEV, and funds can be requested free from the Oasis Devnet Faucet [33]. It is important to note that at the time of this writing, obtaining funds for use within the Oasis Devnet requires identifying information from the user; however, it is assumed that this is only during the early stages of development of the Oasis platform.

Onboarding and Usability

Most users are unfamiliar with activities required for DApp interaction, such as wallet creation, private key storage, and obtaining tokens. As DApps and smart contracts are built on blockchain technology, this requires user actions that are not required by traditional client-server architecture. As the technology matures, we anticipate these activities will become more streamlined, similar to how these functions and interactions have matured on the World Wide Web. Blockchain DApp transactions can appear quite slow in comparison with traditional client-server architecture, but this is an issue that is actively being worked on by the blockchain community.

Malicious Users and Decentralized App Design

DApp users are identified by their wallet's unique public key. Public blockchains, on which most DApps are built, are typically not permissioned, and so anyone can create an infinite number of wallets and unique public keys. As users are able to remain pseudonymous and hide their true identities by proxy of their unique public key, this makes it difficult to ban or blacklist malicious users from accessing the DApp. Similarly, without being able to verify a user's true identity, a malicious user is able to falsely act as another type of user. For example, a malicious participant could pose as a third party or vice versa.

This DApp was designed to allow third parties to continue to add new locations that match the predefined category types into the future. However, these new locations may conflict with previous locations, especially in very dense areas. As the latitude and longitude are rounded to the nearest 4th degree of precision, this results in an accuracy of approximately 11.132 meters with an error of half that distance. Very dense population areas, such

as cities, may have multiple categories of location for that particular area. For example, a gym may exist within less than 11.132 meters of a hospital resulting in more than just one possible category per location. To address this, the DApp was designed to allow future posted categories to overwrite previous ones. However, a malicious user acting as a third party can post an incorrect category for a given location and would be able to overwrite a previously posted category. This would allow the malicious user to iterate through a given location area, post a particular category, and then identify the participants' previous location categories that have changed, which implies that the particular location is the same as that location just posted by the malicious third party. However, because of the transaction costs associated with posting new categories, this iterative process could become prohibitively expensive and thus, may disincentivize this activity. There are also several ongoing blockchain projects working towards the verification of location data in an attempt to combat malicious users [40,41].

Future Improvements

Informed Consent

User onboarding and informed consent are tasks that would require careful design tailored to the particular study and data that are being collected and are out of scope for this tutorial. An informed consent would explain to users that only features of their data were being collected (eg, location category from geocoordinate data) and would take into consideration the legal aspects related to the collection and ownership of data. Although the consent could be managed outside of DApp, a better user experience would be to include a digital informed consent screen within the app, which could also illustrate how the data are used, which parties are involved, and the ability to revoke data sharing.

Expanded Data Features

This DApp provides a basic example of feature extraction for the category of the location from the raw GPS data. One could easily envision more advanced feature extraction for this use case using geolocation data. Additional metadata such as the total distance traveled in a day, variance in the number of locations visited, or the travel radius of participants could be computed by the smart contract, if the smart contract were redesigned to securely store the raw geocoordinates of the locations.

Additional Data Sharing Use Cases

As the source code of the smart contract is publicly available, the smart contract could easily be adapted and shared to fit a variety of needs beyond geolocation data. Other data commonly stored on iOS devices, such as heart rate and steps walked within Apple's HealthKit, could be posted to the smart contract and could make features of the data viewable to interested third parties. Moreover, various data streams could be combined and used within the same smart contract to output additional features, all while safeguarding the sensitive raw data used to create them.

This software has been made publicly available on GitHub at HD2i/GeolocationSmartContract and HD2i/Geolocation-iOS. It was the intention of the authors that improvements, via forks or pull requests, would be made to improve this tutorial. Our goal with this tutorial is to inspire new and unforeseen improvements that would help advance the community as a whole [11,12].

Through this simple use case we have intended to highlight the potential of privacy preservation using smart contracts on blockchain networks such as Oasis. It also illustrates how using a native mobile DApp, privacy-preserving smart contracts could be used to ensure the confidentiality of sensitive health data and at the same time provide researchers with the feature-rich data embedded within.

Acknowledgments

This work was supported by the Institute for Next Generation Healthcare at the Icahn School of medicine at Mount Sinai, and a gift from the Harris Family Charitable Foundation (to JT Dudley).

Conflicts of Interest

None declared.

Multimedia Appendix 1
Demo video.
[MOV File , 27516 KB - jmir_v21i10e13601_app1.mov]

Multimedia Appendix 2
Terms and definitions.
[PDF File (Adobe PDF File), 35 KB - jmir_v21i10e13601_app2.pdf]

Multimedia Appendix 3
Smart contract state variables.
[PDF File (Adobe PDF File), 33 KB - jmir_v21i10e13601_app3.pdf]

Multimedia Appendix 4
Smart contract functions.
[PDF File (Adobe PDF File), 42 KB - jmir_v21i10e13601_app4.pdf]

Multimedia Appendix 5
Steps to deploy geolocation smart contract.
[PDF File (Adobe PDF File), 66 KB - jmir_v21i10e13601_app5.pdf]

Multimedia Appendix 6
Design considerations.
[PDF File (Adobe PDF File), 42 KB - jmir_v21i10e13601_app6.pdf]

References

1. Ethereum Homestead 0.1 Documentation. Dapps URL: http://ethdocs.org/en/latest/contracts-and-transactions/developer-tools.html [accessed 2019-02-01] [WebCite Cache ID 75s1dBrVh]
2. Coravos A, Zimmerman N. Stat. 2018. Blockchains for Biomedicine and Health Care Are Coming. Buyer: Be Informed URL: https://www.statnews.com/2018/07/25/blockchains-biomedicine-health-care-buyer-be-informed/ [accessed 2019-05-01] [WebCite Cache ID 783C3E14r]
3. Biomedical Blockchain Initiative Index. 2019. Healthcare and Biomedical Blockchains: Project Landscape URL: https://db.biomedicalblockchain.org/ [accessed 2019-01-31] [WebCite Cache ID 75qSnenbY]
4. State of the State of the DApps. 2019. Ranking the Best DApps of Ethereum, EOS & Steem URL: https://www.stateofthedapps.com/rankings [accessed 2019-02-01] [WebCite Cache ID 75s6JIGMJ]
5. Statista. 2019. Number of Daily Active Facebook Users Worldwide as of 2nd Quarter 2019 (in Millions) URL: https://www.statista.com/statistics/346167/facebook-global-dau [accessed 2019-01-31] [WebCite Cache ID 75sPvmW48]
6. Griggs KN, Ossipova O, Kohlios CP, Baccarini AN, Howson EA, Hayajneh T. Healthcare blockchain system using smart contracts for secure automated remote patient monitoring. J Med Syst 2018 Jun 6;42(7):130. [doi: 10.1007/s10916-018-0982-x] [Medline: 29876661]
7. Singhal B, Dhameja G, Panda PS. Beginning Blockchain: A Beginner's Guide to Building Blockchain Solutions. New York City: Apress; 2018.
8. GitHub. 2018. Angular-seed — the Seed for AngularJS Apps URL: https://github.com/angular/angular-seed [accessed 2019-02-01] [WebCite Cache ID 75s6tBxXt]
9. Bootstrap. 2019. URL: https://getbootstrap.com/ [accessed 2019-02-01] [WebCite Cache ID 75s73ZbRR]
10. Truffle Suite: Sweet Tools for Smart Contracts. 2019. URL: https://truffleframework.com/ [accessed 2019-02-01] [WebCite Cache ID 75s7AD85I]
11. GitHub. 2019. HD2i: Geolocation-iOS URL: https://github.com/HD2i/Geolocation-iOS [accessed 2019-02-01] [WebCite Cache ID 75s534uIq]
12. GitHub. 2019. HD2i: GeolocationSmartContract URL: https://github.com/HD2i/GeolocationSmartContract [accessed 2019-02-01] [WebCite Cache ID 75s4vJTip]
13. BlockchainHub. 2019. Decentralized Applications – dApps URL: https://blockchainhub.net/decentralized-applications-dapps/ [accessed 2019-05-09] [WebCite Cache ID 78ElGp27o]
14. Altman M, Wood A, O'Brien DR, Gasser U. Practical approaches to big data privacy over time. Int Data Priv Law 2018 Mar 12;8(1):29-51 [FREE Full text] [doi: 10.1093/idpl/ipx027]
15. Swan M. Blockchain: Blueprint for a New Economy. First Edition. Sebastopol, CA: O'Reilly Media; 2015.
16. Jones M, Johnson M, Shervey M, Dudley JT, Zimmerman N. Privacy-preserving methods for feature engineering using blockchain: review, evaluation, and proof of concept. J Med Internet Res 2019 Aug 14;21(8):e13600 [FREE Full text] [doi: 10.2196/13600] [Medline: 31414666]
17. Saeb S, Zhang M, Karr CJ, Schueller SM, Corden ME, Kording KP, et al. Mobile phone sensor correlates of depressive symptom severity in daily-life behavior: an exploratory study. J Med Internet Res 2015 Jul 15;17(7):e175 [FREE Full text] [doi: 10.2196/jmir.4273] [Medline: 26180009]
18. Sano A, Picard RW. Stress Recognition Using Wearable Sensors and Mobile Phones. In: Humaine Association Conference on Affective Computing and Intelligent Interaction. 2013 Presented at: ACII'13; September 2-5, 2013; Geneva, Switzerland p. 671-676. [doi: 10.1109/ACII.2013.117]
19. Müller AM, Blandford A, Yardley L. The conceptualization of a just-in-time adaptive intervention (JITAI) for the reduction of sedentary behavior in older adults. Mhealth 2017;3:37 [FREE Full text] [doi: 10.21037/mhealth.2017.08.05] [Medline: 29184889]
20. Bertz JW, Epstein DH, Preston KL. Combining ecological momentary assessment with objective, ambulatory measures of behavior and physiology in substance-use research. Addict Behav 2018 Aug;83:5-17 [FREE Full text] [doi: 10.1016/j.addbeh.2017.11.027] [Medline: 29174666]

21. Madden M. Pew Research Center. 2014. Public Perceptions of Privacy and Security in the Post-Snowden Era URL: https://www.pewinternet.org/2014/11/12/public-privacy-perceptions/ [accessed 2019-02-01] [WebCite Cache ID 75qnMH4B6]
22. Oasis Dev Docs. 2019. URL: https://docs.oasiscloud.io/en/latest/quickstart-guide/ [accessed 2019-02-01] [WebCite Cache ID 75s2Ga8WQ]
23. Kuo TT, Kim HE, Ohno-Machado L. Blockchain distributed ledger technologies for biomedical and health care applications. J Am Med Inform Assoc 2017 Nov 1;24(6):1211-1220 [FREE Full text] [doi: 10.1093/jamia/ocx068] [Medline: 29016974]
24. Angraal S, Krumholz HM, Schulz WL. Blockchain technology: applications in health care. Circ Cardiovasc Qual Outcomes 2017 Sep;10(9):e003800. [doi: 10.1161/CIRCOUTCOMES.117.003800] [Medline: 28912202]
25. Cheng R, Zhang F, Kos J, He W, Hynes N, Johnson N, et al. Ekiden: A Platform for Confidentiality-Preserving, Trustworthy, and Performant Smart Contracts. In: Proceedings of the European Symposium on Security and Privacy. 2019 Presented at: EuroS&P'19; June 17-19, 2019; Stockholm, Sweden. [doi: 10.1109/eurosp.2019.00023]
26. Canetti R. Universally Composable Security: A New Paradigm for Cryptographic Protocols. In: Proceedings 42nd IEEE Symposium on Foundations of Computer Science.: IEEE; 2001 Presented at: FOCS'01; October 8-11, 2001; Newport Beach, CA, USA URL: https://eprint.iacr.org/2000/067 [doi: 10.1109/sfcs.2001.959888]
27. Ethereum Community. Guia Homestead de Ethereum - Read the Docs. 2018. Contracts URL: http://ethdocs.org/en/latest/contracts-and-transactions/contracts.html [accessed 2019-02-01] [WebCite Cache ID 75s27lzpk]
28. Oasis Dev Docs. 2018. Introducing the Oasis Devnet URL: https://medium.com/oasislabs/oasis-devnet-launch-d580f9e4c800 [accessed 2019-02-01] [WebCite Cache ID 75s33RKv6]
29. GitHub. 2019. Matter-Labs: Web3Swift URL: https://github.com/matterinc/web3swift [accessed 2019-02-01] [WebCite Cache ID 75s2SuA03]
30. Remix - Ethereum IDE. URL: https://remix.ethereum.org [accessed 2019-02-01] [WebCite Cache ID 75s3C1GVd]
31. Remix - Ethereum IDE. 2018. Running Transactions URL: https://remix.readthedocs.io/en/latest/run_tab.html [accessed 2019-02-01] [WebCite Cache ID 75s3MYZgE]
32. MetaMask. 2017. URL: https://metamask.io/ [accessed 2019-02-01] [WebCite Cache ID 75s2uRWC2]
33. Oasis Dev Docs. 2018. Oasis Devnet: Faucet URL: https://faucet.oasiscloud.io/ [accessed 2019-02-01] [WebCite Cache ID 75s2gsKpd]
34. Oasis Dev Docs. 2019. Oasis Contract Kit - Oasis Documentation URL: https://docs.oasiscloud.io/en/latest/contract-kit/ [accessed 2019-02-01] [WebCite Cache ID 75s3fDR2x]
35. Oasis Dev Docs. Web3cjs - Oasis Documentation URL: https://docs.oasiscloud.io/en/latest/web3c.js/ [accessed 2019-06-27] [WebCite Cache ID 75s4IMOwT]
36. GitHub. 2018. Peepeth Client (Unofficial) URL: https://github.com/matterinc/PeepethClient [accessed 2019-02-01] [WebCite Cache ID 75s2Zz3N7]
37. Zang H, Bolot J. Anonymization of Location Data Does Not Work: a Large-Scale Measurement Study. In: Proceedings of the 17th Annual International Conference on Mobile Computing and Networking. 2011 Presented at: MobiCom'11; September 19-23, 2011; Las Vegas, Nevada, USA p. 145-156. [doi: 10.1145/2030613.2030630]
38. Freudiger J, Shokri R, Hubaux JP. Evaluating the Privacy Risk of Location-Based Services. In: Proceedings of the International Conference on Financial Cryptography and Data Security. 2012 Presented at: FC'11; February 28-March 4, 2011; Gros Islet, Saint Lucia p. 31-46. [doi: 10.1007/978-3-642-27576-0_3]
39. Oasis Dev Docs. 2019. Developing Confidential Smart Contracts - Oasis Documentation URL: https://docs.oasiscloud.io/en/latest/confidentiality-develop/ [accessed 2019-02-01] [WebCite Cache ID 75s4IMOwT]
40. Trouw A, Levin M, Scheper S. The XY Oracle Network - White Paper - XYO Network. 2018. The XY Oracle Network: The Proof-of-Origin Based Cryptographic Location Network URL: https://docs.xyo.network/XYO-White-Paper.pdf [accessed 2019-02-01] [WebCite Cache ID 75s45lstf]
41. Foam Space. 2018. FOAM Whitepaper URL: https://www.foam.space/publicAssets/FOAM_Whitepaper.pdf [accessed 2019-02-01] [WebCite Cache ID 75s4C7tCr]

Abbreviations

bpm: beats per minute
DApp: decentralized app
GPS: global positioning system
HD: hierarchical deterministic
ID: identifier
IDE: integrated development environment
iOS: iPhone operating system
JSON: JavaScript object notation
RPC: remote procedure call
TEE: trusted execution environment
VM: virtual machine

Edited by P Zhang, K Clauson; submitted 02.02.19; peer-reviewed by A Corovos, M Hölbl, R Hylock; comments to author 25.04.19; revised version received 14.05.19; accepted 06.08.19; published 23.10.19.

Please cite as:
Johnson M, Jones M, Shervey M, Dudley JT, Zimmerman N
Building a Secure Biomedical Data Sharing Decentralized App (DApp): Tutorial
J Med Internet Res 2019;21(10):e13601
URL: https://www.jmir.org/2019/10/e13601
doi:10.2196/13601
PMID:31647475

Viewpoint

A Research Roadmap: Connected Health as an Enabler of Cancer Patient Support

Gabriel Ruiz Signorelli[1,2,3], MSc; Fedor Lehocki[4,5], MPH, PhD; Matilde Mora Fernández[2], PhD; Gillian O'Neill[3], MSc; Dominic O'Connor[3], MSc; Louise Brennan[3,6], BSc; Francisco Monteiro-Guerra[3,7], MSc; Alejandro Rivero-Rodriguez[7], PhD; Santiago Hors-Fraile[7,8,9], MSc; Juan Munoz-Penas[10], MSc; Mercè Bonjorn Dalmau[10], BSc; Jorge Mota[11], PhD; Ricardo B Oliveira[12], PhD; Bela Mrinakova[13], MPH, PhD, MD; Silvia Putekova[14], PhD; Naiara Muro[15,16,17], MSc; Francisco Zambrana[18], MD; Juan M Garcia-Gomez[19], PhD

[1]Oncoavanze, Seville, Spain
[2]Sport & Society Research Group, Faculty of Educational Sciences, University of Seville, Seville, Spain
[3]Insight Centre for Data Analytics, O'Brien Centre for Science, University College Dublin, Belfield Campus, Dublin, Ireland
[4]Slovak University of Technology in Bratislava, Bratislava, Slovakia
[5]National Centre of Telemedicine Services, Bratislava, Slovakia
[6]Beacon Hospital, Dublin, Ireland
[7]Salumedia Tecnologías, Seville, Spain
[8]Maastricht University, Maastricht, Netherlands
[9]Architecture and Computer Technology Department, University of Seville, Seville, Spain
[10]University of Southern Denmark, Kolding, Denmark
[11]Research Centre in Physical Activity, Health and Leisure, Faculty of Sport, University of Porto, Porto, Portugal
[12]Laboratory of Active Living, Institute of Physical Education and Sports, University of Rio de Janeiro State, Rio de Janeiro, Brazil
[13]First Department of Oncology, Comenius University, Bratislava, Slovakia
[14]Faculty of Health Care and Social Work, University of Trnava, Trnava, Slovakia
[15]Laboratoire d'informatique médicale et d'ingénierie des connaissances en e-Santé, Sorbonne Universités, Paris, France
[16]eHealth and Biomedical Applications, Vicomtech, Donostia-San Sebastian, Spain
[17]Biodonostia, Donostia-San Sebastián, Spain
[18]Department of Oncology, Infanta Sofia University Hospital, Madrid, Spain
[19]Biomedical Data Science Lab, The Institute of Information and Communication Technologies, Universitat Politecnica de Valencia, Valencia, Spain

Corresponding Author:
Gabriel Ruiz Signorelli, MSc
Oncoavanze
Calle Presidente Cardenas, 8 Local Oncoavanze
Seville, 41013
Spain
Phone: 34 954 231 459
Email: gasignorelli@gmail.com

Abstract

The evidence that quality of life is a positive variable for the survival of cancer patients has prompted the interest of the health and pharmaceutical industry in considering that variable as a final clinical outcome. Sustained improvements in cancer care in recent years have resulted in increased numbers of people living with and beyond cancer, with increased attention being placed on improving quality of life for those individuals. Connected Health provides the foundations for the transformation of cancer care into a patient-centric model, focused on providing fully connected, personalized support and therapy for the unique needs of each patient. Connected Health creates an opportunity to overcome barriers to health care support among patients diagnosed with chronic conditions. This paper provides an overview of important areas for the foundations of the creation of a new Connected Health paradigm in cancer care. Here we discuss the capabilities of mobile and wearable technologies; we also discuss pervasive and persuasive strategies and device systems to provide multidisciplinary and inclusive approaches for cancer patients for mental well-being, physical activity promotion, and rehabilitation. Several examples already show that there is enthusiasm in strengthening

the possibilities offered by Connected Health in persuasive and pervasive technology in cancer care. Developments harnessing the Internet of Things, personalization, patient-centered design, and artificial intelligence help to monitor and assess the health status of cancer patients. Furthermore, this paper analyses the data infrastructure ecosystem for Connected Health and its semantic interoperability with the Connected Health economy ecosystem and its associated barriers. Interoperability is essential when developing Connected Health solutions that integrate with health systems and electronic health records. Given the exponential business growth of the Connected Health economy, there is an urgent need to develop mHealth (mobile health) exponentially, making it both an attractive and challenging market. In conclusion, there is a need for user-centered and multidisciplinary standards of practice to the design, development, evaluation, and implementation of Connected Health interventions in cancer care to ensure their acceptability, practicality, feasibility, effectiveness, affordability, safety, and equity.

(J Med Internet Res 2019;21(10):e14360) doi:10.2196/14360

KEYWORDS

cancer; Connected Health; mHealth; eHealth; mental health; physical activity; rehabilitation; wearable; Internet of Things; quality of life

Introduction

Improvements in cancer diagnosis and treatment have resulted in increased survival rates of those suffering from cancer. Some cancer types can now be considered chronic diseases and the attention to patients' quality of life (QoL) during and after cancer treatment is increasing [1]. Consequently, there is an increased burden to the cancer care system due to the increased level of complexity. A clear example of the increased level of treatment complexity in cancer care is the emergence of precision medicine in oncology [2], which relies heavily on the use of biomedical data-driven approaches for choosing the right therapy at the individual and molecular level [3].

According to the World Health Organization, the achievement of universal health coverage to ensure healthy lives and to promote well-being for all is only possible by using new technologies [4]. The advent of Connected Health creates an opportunity to overcome barriers to health care among patients diagnosed with cancer. The technology industry has been developing evidence-based technologies to improve physical activity programs in the clinical setting. The majority of adults in developed countries own mobile phones, therefore, most of the mobile phone-based medical apps aimed at enhancing patient care are easily accessible [5].

Connected Health is defined as follows [6]:

> *Connected Health encompasses terms such as wireless, digital, electronic, mobile, and telehealth and refers to a conceptual model for health management where devices, services, or interventions are designed around the patient's needs, and health-related data is shared, in such a way that the patient can receive care in the most proactive and efficient manner possible. All stakeholders in the process are connected by means of timely sharing and presentation of accurate and pertinent information regarding patient status through smarter use of data, devices, communication platforms, and people.*

Consequently, Connected Health is about not only technology but also the transformation of health care. Thus, it can provide the foundation for the transformation of cancer care into a patient-centric model focused on providing personalized support to the specific needs of each patient.

In recent years, Connected Health technologies have been applied to cancer care in multiple settings, including screening, patient treatment, after-treatment management, and follow-up of survivors. Connected Health has been promising in managing the quality of care of cancer survivors; this is because it allows the patients to communicate with the health care providers regarding their health status without any dependency on the capacity of a physician [7]. Connected Health allows health care professionals to collect real-time data, which allows medical needs to be identified and responded to within appropriate time frames [7]. Connected Health technologies assist in the evaluation of treatment efficacy by monitoring complications of treatments, such as toxicities or adverse events, and by collecting patient-reported outcome measures (PROMs). Some standard scales and terminologies report these outcomes in a standardized way to be used by digital systems, such as the International Consortium for Health Outcomes Measurement (ICHOM) [8] and the Common Terminology Criteria for Adverse Events (CTCAE) [9]. In addition, new research has shown the importance of PROMs for building integrated cancer care [10]. Limitations to the current application of Connected Health in oncology are the lack of regulation of such technologies and their susceptibility to cybercrime [7].

In this viewpoint paper, we will provide an overview of important areas for the creation of a new Connected Health paradigm in cancer care.

Cancer Patient Care

Mental Well-Being

Alongside the physical symptoms associated with cancer and its treatment, people affected by cancer may also experience mental challenges, such as feelings of uncertainty, fear, sadness, or distress [11]. While these symptoms are often transient, at least 15%-20% of cancer survivors will experience disabling levels of distress, such as anxiety or depression, as late as 10 years after their initial diagnosis [12]. Compared to the general population, there is also a greater risk of suicide in cancer survivors [13,14]. Psychological distress can also interfere with one's cognitive, emotional, and behavioral resources to

effectively cope with the impact of living with a diagnosis of cancer [15]. It has been associated with impaired QoL and poor treatment adherence [11]. Psychological distress can also act as a barrier to engaging in recommended long-term self-care activities such as physical activity [16]. Helping cancer survivors better cope with the consequences of cancer and its treatment can enhance their mental well-being, prevent long-term distress, and increase engagement in recommended, positive, health-related behaviors for a better QoL.

Psychosocial interventions, such as cognitive behavioral therapy and behavior change interventions, can enhance coping skills in people affected by cancer [17]. However, face-to-face psychological treatment is costly, often stigmatized, and not always available in hard-to-reach areas. This can result in cancer survivors not having timely access to the psychological support they need [18]. Technological advances in health care and Connected Health, in particular, provide opportunities for increasing access to supportive and individually tailored psychological support for those who need it and can be seen as a first step to seeking support [19]. In supporting mental well-being in people affected by cancer, Connected Health interventions have, to date, focused on the development of Web-based delivery of evidence-based therapies, videoconferencing with health care professionals, and websites for self-care interventions such as coping-skills training [20,21]. While most Connected Health interventions focus on the use of technologies already available and in use in a person's day-to-day life, gaming technologies and wearables are being adapted for use in supporting mental well-being across age groups [21].

Connected Health interventions for mental well-being in people affected by cancer are generally found to be acceptable; users report high levels of satisfaction with interventions that contain elements of social support, access to health care professionals, and tracking through easy-to-interpret visual representations of measured outcomes [21]. Despite this, the evidence for improvements in psychosocial outcomes remains mixed. While such interventions have been found to be effective in enhancing outcomes such as self-efficacy, coping, and perceived social support, there is limited evidence for their impact on more severe symptoms of distress, such as anxiety or depression [20,21]. These interventions have largely been evaluated in mixed cancer populations, suggesting that they may not cater sufficiently to individual psychosocial needs [22]. Mixed findings can also be explained by the heterogeneity in the technological tools that are used, their exact content, and their duration and intensity of use, as well as in the differences in methods employed for designing, developing, and evaluating these eHealth interventions [23]. While there is promise for the use of eHealth in supporting mental well-being in people affected by cancer, a standardization of practices is needed to ensure that evidence-based technology solutions made available to potential users are effective as well as acceptable, practical, safe, equitable, and affordable [21,23].

Across the development of Connected Health interventions for mental health more broadly, there has been an urgent call for an internationally and cross-disciplinary set of agreed principles and practices for the research and evaluation of digital tools surrounding the intervention's effectiveness, user experience, and adherence, as well as data safety and privacy [24]. These are equally relevant areas to the design and implementation of Connected Health tools for mental well-being within the cancer population. The effectiveness of an intervention depends upon the correct application of relevant theories and behavior change techniques. However, many Connected Health solutions for mental well-being in people affected by cancer do not provide evidence of this process [21,23,25]. With advances in big data analytics and machine learning algorithms, future developments in this area have the potential to enhance our understanding of behavior change by identifying the digital health strategies that will be most effective in supporting mental well-being based on individual characteristics; this will allow for a more effective and personalized intervention [23]. We need to go beyond effectiveness and promotion of engagement and long-term adherence to the use of such technologies by people affected by cancer; to this end, it is recommended that user-centered design processes, including the involvement of a variety of identified stakeholders and a multidisciplinary team, also be applied from the point of design through evaluation and implementation [23-25].

Advances in technology have enabled new ways to both passively and actively collect user data through methods such as global positioning system, photos, and voice-recording. This highlights the need for guidelines to be developed to ensure potential users have a clear understanding surrounding the storage, use, and sharing of their data [24]. However, this also raises issues with data privacy and safety in the field of mental health care. Torous et al recommend that as technologies advance in this field, there will be a need to undertake technical reviews and audits [24]. Finally, specific regulations and safeguards also need to be implemented to ensure that potential users can objectively assess the quality of Connected Health tools and select the most appropriate tool matching their needs. Evidence-based Connected Health tools for mental health need to be regularly revised and re-evaluated to account for the latest evidence-based knowledge in the area. The application of an evidence-based treatment to Connected Health does not mean that the technology itself is evidence based; it is essential that the technology undergoes its own clinical evaluation [24].

Physical Activity and Rehabilitation

Physical exercise can be challenging for many patients with cancer due to disease and treatment side effects. Physical rehabilitation in cancer, delivered by health care professionals such as physical therapists and exercise physiologists, assists individuals who have experienced disability in achieving and maintaining optimal functioning in interactions with their environments [26]. Several studies support the integration of physical activity and exercise into ongoing treatment and management across the cancer trajectory [27-29]. Cancer survivors are advised to meet the recommended physical activity levels for adults of at least 150 minutes per week of moderate-to-intense aerobic exercise and muscle-strengthening activities at least two days a week [27,30,31]. A holistic approach that considers disease-related factors, such as severity of the disease, symptoms, and treatment status, and addresses modifiable barriers is essential to facilitate the uptake of physical

activity in this population [32-34]. As such, interventions that provide additional support during physical rehabilitation should be considered.

Mobile apps and wearable sensor systems have the potential to support patients and assist in delivering high-quality physical rehabilitation when used as adjuncts to professional treatment. Wearable devices are the most ubiquitous technologies used to track physical activity. As sensors that measure a variety of biological and physiological parameters become smaller and less costly, they can be integrated into wearable systems, which allow for real-time monitoring and evaluation of rehabilitation in a discreet, cohesive manner [35]. When combined with an Internet-enabled interface (eg, a mobile app) and advanced data analytics, these systems can collect and process data and provide personalized feedback to users and to health care providers. Wearable technology is currently used to measure and analyze a wide range of rehabilitation activities and effects; these include movement and physiological responses to exercise as well as health factors closely associated with rehabilitation, such as sleep and general physical activity [35,36]. These devices provide real-time feedback and allow health care professionals to remotely monitor clinically relevant variables in a nonclinical setting. While there are opportunities for consumer-grade wearable systems in oncology practice [37], in comparison to bespoke rehabilitation systems, these lack clinical validity and are not equipped to conduct human motion analysis at the level often required in a rehabilitation setting [38,39].

Wearable technologies can be used to deliver exercise to those who may experience exercise-limiting symptoms. Wearable neuromuscular electrical stimulation (NMES) technology may be a pragmatic alternative to voluntary aerobic and resistance exercise in such cases. NMES involves the contraction of skeletal muscles via electrical impulses delivered to motor nerves using surface electrodes placed over target muscle groups [40]. Traditional high-frequency NMES (20-100Hz) has been extensively used in sports training and in different rehabilitation contexts to augment muscle mass and strength [41]. In addition, emerging evidence demonstrates the efficacy of low-frequency NMES (3-12 Hz) for enhancing cardiorespiratory fitness [42]. Furthermore, wearable NMES technology allows for the delivery of NMES exercise sessions safely and unsupervised in the user's home; this makes it an attractive technology for patients with cancer who may be physically deconditioned, with impaired cardiorespiratory and muscular fitness. Previous work has attempted to implement NMES exercise into the rehabilitation of advanced cancer patients. The NMES exercise interventions used were adapted from orthopedic and neurological rehabilitation contexts and were largely unsuccessful due to inappropriately designed protocols [43,44]. Recent work using a personalized and progressive NMES exercise approach, designed with early-stage cancer rehabilitation in mind, has reported improvements in functional strength and QoL outcomes [45]. This preliminary evidence highlights the potential of wearable systems such as NMES exercise technology in cancer rehabilitation.

Providing patients with additional support during exercising may help reduce barriers to home-based rehabilitation. One such digital rehabilitation technology is biofeedback, which is the process of providing an individual with information regarding a particular body function and allows the individual to self-regulate this function. Through visual, audio, haptic, or multi-model feedback delivered in real time or after exercising, patients receive personalized guidance on their rehabilitation. This may help enhance motor learning and improve engagement with therapy [46,47]. Biofeedback systems in cancer rehabilitation should be codesigned with users for specific clinical contexts to ensure that they are meaningful, accessible, and clinically effective [48,49]. Including appropriate elements of gamification (eg, points and levels) in the software development process may help promote engagement and adherence in some cohorts [50,51].

Gamification helps make everyday tasks more interesting and, thus, potentially more engaging for the users [52]. The health care industry continuously embraces this idea of gamification as a way of encouraging patients and individuals diagnosed with chronic diseases to enjoy physical activities. Gamification refers to the process of using the principles found in game design in a nongame context. Studies reveal that gamification technology in health care is the best way of engaging people in physical activities [53]. However, it is important to understand that the patient's motivation to engage in physical activities through mobile technology may be influenced by social interaction, attitude, and individual personality [53,54]. Therefore, health care practitioners may need to consider goal-oriented activities that will help patients understand the primary objectives of the games on the mobile technology along with aspects that will influence patients' engagement with the games.

Promoting physical activity among cancer survivors remains a significant priority in health care. Health care practitioners are continuously looking for ways they can enhance patient engagement in physical activities [55]. The introduction of Connected Health presents limitless possibilities because it strengthens the patient-practitioner relationship. Connected Health allows coaching of physical activity because it enables exchange of valuable information between the patient and health care practitioner that can help in managing the patient's physical activity [56]. In a recent study, a continuous wrist-mounted tracker monitored physical activity with daily step counts in 38 patients with different cancer types undergoing chemoradiotherapy with curative intent; physical activity monitored by the tracker correlated significantly with the risk of unplanned hospitalization during or shortly after therapy, which suggested a role for this technology as a dynamic indicator of this potentially life-threatening adverse event [57].

Additionally, clinicians may benefit from having quantifiable, reliable, and objective measures of therapeutic assessments and interventions that were previously not possible [58]. They may also benefit from high-volume, longitudinal, granular data, with which they can analyze disease management and demonstrate outcomes to commissioners. Researchers using these systems will gain the ability to collect large-scale, momentary, longitudinal data; quantify intervention outcomes; and supplement existing QoL assessment tools.

Persuasive and Pervasive Technology in Cancer Care

Internet of Things

The use of pervasive computing solutions, such as mobile phones, wearables, or connected home appliances, can help monitor and assess the health status of patients. These environments of connected devices can be described within the concept of *Internet of Things*, which was defined by Gartner as "the network of physical objects that contain embedded technology to communicate and sense or interact with their internal states or the external environment" [59]. These connected technologies provide multiple ways to monitor the real-time health status of patients by sensing the individual and his or her environment [60]. More advanced systems that can be worn by patients include textile wearable devices; flexible, stretchable, and printable devices (eg, epidermal electronics and e-skin); and devices embedded into a patient's living environment.

It has been shown that even the simplest pervasive technology, such as regular phone calls and mobile phone apps, used for structured follow-up intervention provides improvements in pain, depression, distress, and QoL. These studies were focused on patients after completion of their primary treatment of breast, prostate, colorectal, and lung cancer [61].

For purposes of cancer care, sensing systems can be used for objective older adult patient monitoring for baseline functional status and treatment toxicity [62]. For immobile patients, the big challenge is prevention of pressure ulcers. Wearable technology can support health care providers in keeping turning schedules for large groups of patients [63]. Skin conductance or temperature can provide important insights into physiologic changes in cancer patients. Textile wearable technology integrates sensor, energy sources, processing, and communication devices within the garment. An example of such a system is a bra that helps to detect breast cancer using intelligent breast patches that detect small circadian temperature changes in breast cells and communicate this via mobile phone to a designated remote health center [64]. Electronic circuits in flexible electronics are manufactured or printed on flexible substrates (eg, cloth fabrics and the human body). It can be used to detect specific biomarkers, such as complement proteins associated with breast cancer from saliva, tears, or the breath [65]. Epidermal electronics, with their stretchability, enable devices to bend to a very small radius and still maintain the desired electronic performance. This allows the design of an ultra-thin surface electromyography patch for swallowing-therapy exercises for head and neck cancer patients [66]. The photoplethysmographic sensing method enables assessment of cuffless blood pressure using the integration of electrodes into chair pads and arms or into beds under the mattress [67].

A cognitive perspective presented in Lucchiarri et al [68] is related to methods for nutrition monitoring in cancer patients. Inspiration for food classification and estimation (eg, macronutrient content) can be found from other chronic diseases like diabetes [69]. The solution presented in this paper was developed by a research team at the National Technical University of Athens and is based on the classification system for food images taken by a mobile phone camera and exploiting convolutional neural networks.

Personalization

An underlying challenge of mobile health (mHealth) apps is user abandonment [70-72]. This problem has also been reported in studies with cancer survivors, with the majority of users ceasing app usage within 1-3 months [73,74]. Among the factors associated with this lack of long-term engagement is the feeling of low perceived personal relevance in the experience provided by these apps. It is believed that mobile-based interventions that are closely tailored to the individual's convictions and motivations are more likely to be observed and remembered [75]. Therefore, personalization or tailoring can help increase the intended effects of communication, which can contribute to overcoming user abandonment and improving the effectiveness of these systems [76]. This need for tailored experiences has been highlighted in several studies with cancer survivors [77-79]. mHealth apps can address this need as they provide an opportunity to access actionable individual-level data; this facilitates the tailoring of interventions to users' personal information, such as treatment history, age, gender, stage of change, fitness level, cognitive flexibility, or health goals [74]. However, there is a paucity of research exploring this topic. Only recently has there been research on mHealth tools to connect patients with timely and actionable educational health information [78]. Overall, mHealth systems for cancer lack the personalization that could engage users in their long-term adoption.

Cancer survivors reported a preference for an app that is highly individualized to them and suggested that the app should offer the following: content that is sensitive to user-identified information (eg, cancer diagnosis, personal health considerations, age, and physical limitations); a feature to self-monitor changes in health indicators (eg, heart rate, cholesterol, and waist measurement); personalized goal suggestions; personalized role model narratives; suggestions based on users' location and weather; and adaptation to trends over time [79].

The theory and practice in the field of computer-tailored health communication [80-84], particularly the one applied to real-time coaching systems [85], proposes a variety of approaches to adapt the four different properties of communication—timing, intention, content, and representation—to create more individual technological experiences. These approaches include the following:

1. Feedback: presenting the user with information about themselves.
2. Interhuman interaction: support for interaction with other people with a similar condition.
3. Adaptation: direct messages regarding an individual's status on key theoretical determinants.
4. User targeting: increase attention or motivation by conveying that the communication is designed specifically for the user.

5. Goal setting: learn user-specific goals based on individual patterns.
6. Context awareness: tailoring communication based on external information.
7. Self-learning: learning about the user's reactions to previous communications.

Such strategies are aligned with the tailoring needs reported by cancer survivors [77,79,80] and can be used to complement the guidelines for empowering cancer survivors through mobile apps [86,87].

A recent randomized controlled trial tested a technology-enhanced lifestyle program in prostate cancer patients after treatment for localized disease. Prostate cancer patients are at risk of relapse and lifestyle behaviors may reduce this risk. A combination of a website, Fitbit One, and short message service (SMS) text messaging were used to facilitate the adoption of eight behaviors, including vigorous activity, smoking cessation, and dietary improvements. As compared with the control group (32 participants in each arm), significantly more patients in the intervention group modified the score of lifestyle behaviors, mainly driven by diet rather than exercise [88]. This demonstrates feasibility and a high degree of patient acceptability of these interventions that combine education, longitudinal collection of data, and some interface with the Internet or connected devices.

Even though the potential for personalization in mHealth systems seems clear, there are still many barriers that need to be overcome. These include frequent manual input from the user, which can be perceived as burdensome and may decrease interest in app use. Real-time monitoring technologies can partly reduce this entry burden. However, obtaining monitoring data for timely coaching is not as of yet a straightforward process. Most commercially available activity-monitoring devices only allow the capture of data through their Web application programming interfaces, which results in significant delays from the moment the data is collected until the feedback can be provided to the user. An alternative is to collect the raw data from the sensors, which can be done either from these external wearable devices with the appropriate permissions or from the built-in sensors in the device. However, this normally requires complex processing to make sense of such data and involves extra considerations in terms of hardware and software compatibilities.

Currently, there is still little involvement of cancer health care stakeholders in the design and development of these apps. When it comes to mHealth system evaluation, adopting the optimal methodological approaches may be time and resource consuming and, therefore, challenging to put into practice. Hence, there is a paucity of studies evaluating these systems in a structured and controlled manner and assessing the long-term effects [73]. Furthermore, there is a lack of proper evaluation of particular persuasive mechanisms such as those related to providing a personalized experience on user engagement and behavior change, which introduces a challenge regarding the choice of particular strategies for the design of these systems [73].

The integration with electronic health records (EHRs) may significantly help create personalized information for the users; however, this may also be used to generate meaningful health data that can be shared with the professionals. This will reduce manual data entry by users and will increase the value, reliability, and credibility in these systems [89]. Furthermore, developments in mobile and wearable technologies will continue increasing the capability of mHealth apps to monitor health parameters and will bring new opportunities for real-time feedback and motivation. New and more intelligent forms of personalization will be implemented, which can potentially make a difference in maintaining user engagement and increase the intended effects of these systems. Systems will, for example, be able to dynamically adapt content and functionality according to patients' most pressing needs. The *user* can be seen as a dynamic entity; apps will be required to update their internal model of the user by recording and learning from the user's interactions with the app [85]. These apps will then be able to make sense of the user's schedule and routine, progress, and preferences for certain features or particular information. In addition, this information may be used to help create an app experience that is more targeted to the individual needs of the patient.

Personalized mHealth apps can potentially benefit cancer survivors by providing a platform for self-tracking and self-management. This may help gather information on their condition, which may enhance communication with health care professionals and increase self-awareness and self-confidence in achieving health goals. With the rapid rise of mobile phone use and the increase in the complexity and accuracy of mobile monitoring technologies, these apps can now gather large amounts of users' health data and provide information and motivation anytime and anywhere to these individuals. Future research is needed that incorporates theory and practice of computer-tailored health communication in these mHealth systems; this would help leverage individual data and provide highly personalized support and encouragement for self-management in cancer survivorship.

Artificial Intelligence and Machine Learning

The use of digital tools to gather health-related data from patients is widely accepted. However, examining how this data can be applied to enhance and adapt the provision of care to meet individual patients' needs across the health care pathway is still in its early stages. Artificial Intelligence (AI) techniques, based on directly and indirectly reported patient data, have the potential to create personalized supportive recommendations. While the future of AI in cancer care is promising, further research is needed before this type of data-driven solution can become an established method for enhancing patient care.

In the provision of cancer care, the application of AI has been explored as a means to support an earlier and more accurate diagnosis and to improve prognosis [90]. AI aims to go beyond evidence-based medicine by extending the validated and reported knowledge acquired through the discovery of new insights; this will be done by applying data-driven approaches. Hence, it can extend the knowledge derived from model-driven approaches, such as clinical practice guidelines, by providing new patterns and understanding discovered through the analysis of patients' health-related data. One of the most popular

applications of AI is machine learning, which has been used to predict cancer patients' responses to different drug treatments. For example, a recent study [91] showed that AI can predict treatment response, with more than 80% accuracy, by combining learning techniques with extensive patient data. That prediction may help to provide patients with the most effective treatment for them, thereby improving their experience through their patient journey as well as their health outcomes.

However, AI can go beyond enhancing clinical decision support. For instance, a 2018 study evaluated the impact of using AI to optimize the management of cancer-related pain. Through a randomized controlled trial, it was found that AI reduced perceived pain in 20% of patients and decreased inpatient hospital admissions by 40%. Although this solution increased anxiety, it shows that AI has the potential to be applied in more innovative ways within the cancer care pathway [92]. AI can also be used to investigate and analyze cancer patients' behaviors and emotions to generate actionable insights that have the potential to improve patient health-related outcomes. For example, there exists a ready-to-use solution—the recently presented Patient-Reported Information Multidimensional Exploration (PRIME) framework [93]—which uses machine learning and natural language processing on the contents of online patient support groups. This framework is applicable to different cancer types and its code is available online for use [94].

Furthermore, cancer patients' health-related behaviors, such as maintaining a physically active lifestyle, can alleviate cancer-related symptoms including fatigue. AI has previously been applied in other areas of health to influence patients' health-related behaviors, for instance, healthy eating [95] or smoking cessation [96]. However, few of these applications combine AI with validated models of behavior change. Hence, by taking into consideration nuances of cancer patients' needs in combination with behavioral change models, it is possible to create more robust solutions.

The Data Infrastructure Ecosystem

Cooperative care requires operability between all principals in health care, including persons, organizations, devices, applications, and components [97]. The European Commission has officially recognized the need for improving interoperability in Connected Health and has allocated resources to this end. For example, the 2018 eHealth Interoperability Conformity Assessment Scheme for Europe (EURO-CAS) is paving the way for more eHealth interoperability in Europe [98]. This shows evidence for the need for processes and protocols to enable interoperability. The researchers, health care professionals, and other involved parties in the health care ecosystem are well aware of these needs. As such, several organizations for developing standards for health care interoperability have been created, such as Health Level Seven (HL7) International, the International Health Terminology Standards Development Organisation (IHTSDO), the Organization for the Advancement of Structured Information Standards (OASIS), and the Object Management Group (OMG) [97]. They have proposed several approaches for increasing Connected Health interoperability, with initiatives such as Fast Healthcare Interoperability Resources (FHIR), perhaps the most promising up to now for its easy implementation.

Interoperability is key when developing Connected Health solutions because it highly affects integration with health systems and the EHR system [99], thus making these solutions more useful. FHIR came from HL7 [100], which presented shortcomings in the form of needs for complex custom tooling. FHIR aimed to simplify and accelerate HL7 adoption by being easily consumable but robust and by using open Internet standards, where possible, using an easily consumable format. The FHIR approach was based on the representational state transfer (RESTful) principles described by Fielding [101]. FHIR is considered the de facto standard for interoperability in health recommender systems and many solutions are built on top of that [102,103].

The field of PROMs, adverse events, and QoL is of great interest for data collection and further intervention in cancer care. Digital platforms may facilitate self-reporting of signs and symptoms that correlate with other relevant outcomes, such as performance status, treatment adherence, and survival. Data capture in real time of validated measures of PROMs and QoL is facilitated by wearable devices, which have been shown to be equivalent to previous techniques [104]. Interestingly, Web-based symptom monitoring during chemotherapy among breast, genitourinary, gynecologic, and lung cancer patients was associated with improved QoL, fewer visits to the emergency room, and fewer hospitalizations than among the nonintervention group [105]. Furthermore, in a subsequent analysis of overall survival, there was a 5-month, statistically significant difference in survival favoring the intervention group [106].

The Connected Health Economy Ecosystem

Health care systems worldwide currently struggle with challenges of an aging population alongside budgetary pressures. Connected Health is an emerging field with the potential to be a relevant driver in the needed transformation of health care for increasing its quality and efficiency. mHealth allows the implementation of tools to tackle these challenges by improving the efficiency of the health care system and supporting the shift toward prevention [4].

Sensors and mobile apps gather considerable amounts of medical, physiological, lifestyle, daily activity, and environmental data, which could serve as a basis for evidence-driven practice and research activities. mHealth also facilitates patients' access to their data anytime and anywhere and, therefore, facilitates patient empowerment. Hence, it also enhances the connection between the patient and health care professional, thus making it more efficient across contexts. For example, with the help of self-assessment or remote monitoring solutions, patients could manage to live more independently in their home environment. With this additional data, the health care professional may be able to make a more accurate assessment and provide better personalized treatments. In

addition, mobile apps can also be used to encourage adherence to a healthy lifestyle.

The worldwide uptake of Connected Health can be linked to the growth in wireless subscriptions, which has reached over 6 billion internationally [107]. The fast development of communication technologies and health care devices, alongside the social demand, are creating new businesses focusing within highly promising markets. Market potential differs depending on the context of each individual country, thus shaping business. While high-income countries are driven by cutting health care costs and giving patients a more active role in their own health care, developing countries are focusing on improving access to primary care [4].

Approximately 70% of mHealth apps target the consumer wellness and fitness segments. The other 30% target health care professionals by facilitating access to patient data, patient consultation, monitoring, diagnostic imaging, and access to pharmaceutical information, among others [108]. The European Union estimated that there were 97,000 mHealth apps and 3 billion mHealth app downloads in 2015 [109], with these numbers continuing to grow. However, the global market report on mHealth mentions that almost 70% of these apps did not reach sales of US $10,000 per year [109]. The lack of success in monetization of mHealth apps reveals how difficult it is to reach sustainability in this sector. Nevertheless, mHealth business is growing exponentially, from US $5.7 billion in 2015 to an estimation of US $31 billion in 2020, making it both an attractive and challenging market [109].

Digital solutions facilitate complex monetization models where different and complementary models can coexist in a single app, making the market appealing and challenging at the same time. Monetization models can be classified into eight groups:

1. Free: there is no payment for the use of the app.
2. Download: one payment for a license for use.
3. Subscription: a license for a period of time.
4. Platform: income for installation and use in an organization.
5. Service: income for a service, content, and products.
6. Marketing: incomes for advertisements and content placement.
7. Results based: the improvement of the efficiency and/or health of patients.
8. Assets: income based on selling data generated by the use of the app.

In addition, it is possible to have multiple monetization models for a single solution. One app can be used by different stakeholders at the same time in different ways with different needs, which means complementary monetization paths; for example, free for the user but using marketing (eg, advertisements or contents) unless the user pays a premium service.

One of the most challenging monetization models is the building of assets based on data generated by users. mHealth apps can gather large quantities of information, which can be processed for different purposes. Big data analysis is the capacity to analyze large amounts of unstructured datasets from a wide range of sources, involving the extraction of potentially valuable information in a cost-effective way. Therefore, the use of health-related data is a central element of epidemiological research as it can enable the identification of patterns to improve treatments, optimize trial periods for medication, and advance mechanisms for early detection and prevention of diseases. This potential also allows for the development of new, innovative business models in health care.

Connected Health Barriers

The popularization of mobile phones and, more recently, wearable devices is creating a large volume of population data and creates an opportunity for Connected Health in cancer care. It is likely that future patients will generate their own data, either because such devices have become part of the care practice model or because these devices are part of the patients' lives. Consequently, data interoperability problems may become even more relevant than they are today. For instance, one of the most popular wearable devices for tracking health-related information that may contribute to developing AI-based solutions for cancer patients is the Fitbit, which still uses a proprietary architecture. This hampers its usage in AI systems that are compatible with data following existing industry standards, such as HL7 Consolidated Clinical Document Architecture (C-CDA) or the upcoming FHIR.

Nevertheless, the adoption of Connected Health in cancer care still has many barriers to address. These systems need a robust regulatory framework to ensure they are of high quality, they demonstrate clinical effectiveness, and that privacy and security are respected and protected. Existing limitations are related to an inherent digital divide, concerns about privacy, data volume, cost, and reimbursement. It is imperative that these solutions are interoperable and have the potential to be smoothly synchronized with EHRs. Indeed, if siloed, the capacity for Connected Health solutions will be limited regarding their support for cancer patients as part of the digital-based cancer care system to which developing countries are transitioning. In this context, there already exists standardized communication frameworks and data architectures to facilitate such interoperability in the data provided by Connected Health cancer solutions.

Regardless of the data origin, the reasoning behind the recommendations generated by AI-based systems often remains difficult to explain, as these systems can be opaque and nonintuitive. This may result in a lack of trust in AI-based systems, reducing their adoption by patients, and may also prevent health care professionals from prescribing them. Future systems will have to strive to explain how recommendations have been computed [110].

Privacy and Data Protection

The processing of data concerning personal health records is particularly sensitive and requires special protection. Personal data protection is a fundamental right in Europe and, therefore, vital for building trust in mHealth solutions [111].

This is particularly important considering these personalization technologies may deal with sensitive personal data from users.

Even if technologies adopt the appropriate information technology infrastructure for privacy requirements, users can be hesitant to share their personal information. Patients need to think about the trade-off between maintaining privacy of the data collected through these mHealth apps and the potential benefit this might bring in facilitating a more individualized and engaging experience [112]. Patients also need to feel that they can rely on the app; they should also be aware that the information is coming from a trustworthy source (eg, their cancer care providers or an authoritative health agency) or that experts and health care professionals have validated it.

Connected Health solutions should be especially concerned with ensuring data protection and privacy for all individual citizens of the European Union and the European Economic Area. Moreover, given the international component of connected services, addressing the transfer of personal data outside the European Union and the European Economic Area would be of high importance. This may be in compliance of the General Data Protection Regulation [113] by giving control to individuals and by homogenizing the environment for international business activities relative to personal data.

Conclusions

Connected Health in cancer care aims to reduce barriers to health care provision through the development of well-designed, usable, and validated digital solutions. There is an urgent need for user-centered and theory-driven international and multidisciplinary standards of practice to inform the design, development, evaluation, and implementation of Connected Health interventions in cancer care. This can ensure that solutions are seen as acceptable, practical, effective, affordable, safe, and equitable. Including users in the design process, testing extensively, and committing to the development of a regulatory framework for software components are strategies that need to be adopted when developing a Connected Health app for cancer care.

Regarding the technology developments currently available in Connected Health, there is early evidence showing that AI can be used to compute recommendations aiming to support cancer patients. Classical model approaches can then be augmented with the new evidence inferred from data analytics. Increasing sources of patient data can enhance the potential of AI by facilitating opportunities for improved user modeling and evaluation of the actual evidence through this data. The Connected Health market is growing and there are many ways it can be made economically viable. The volume of generated data offers great opportunities to extract actionable information for patients; however, interoperable and comprehensive approaches to their use are crucial. The field of machine learning and advances in big data analysis, in particular, provide opportunities for the development of effective Connected Health interventions in cancer by identifying the content and extent of support needed by a user: that is, being able to identify what intervention works for whom and under what circumstances.

Acknowledgments

This article/publication is based upon work from COST Action ENJECT TD 1405, supported by COST (European Cooperation in Science and Technology; www.cost.eu).

The authors thank Luis Fernández-Luque and Päivi Salminen for their contributions. GRS, FMG, GON, DOC, LB, JMP, and MBD are part of Cancer: Activating Technology for Connected Health (CATCH) and are supported by a grant from the European Union's Horizon 2020 research and innovation program under a Marie Sklodowska-Curie grant (grant number: 722012). FL was supported by grants from the Tatra banka Foundation (2018vs108) and the Research Agency in Slovakia (ITMS: 26240120038). JM was supported by grants from Fundação para a Ciência e a Tecnologia (FCT; grant numbers: SFRH/BSAB/142983/2018 and UID/DTP/00617/2019) as well as from the Programa de Bolsas Santander Universidades 2018.

Conflicts of Interest

None declared.

References

1. van Weert E, Hoekstra-Weebers JE, May AM, Korstjens I, Ros WJ, van der Schans CP. The development of an evidence-based physical self-management rehabilitation programme for cancer survivors. Patient Educ Couns 2008 May;71(2):169-190. [doi: 10.1016/j.pec.2007.11.027] [Medline: 18255249]
2. Shin S, Bode A, Dong Z. Precision medicine: The foundation of future cancer therapeutics. NPJ Precis Oncol 2017;1(1):12 [FREE Full text] [doi: 10.1038/s41698-017-0016-z] [Medline: 29872700]
3. Smith NG, Gyanchandani R, Shah OS, Gurda GT, Lucas PC, Hartmaier RJ, et al. Targeted mutation detection in breast cancer using MammaSeq™. Breast Cancer Res 2019 Feb 08;21(1):22 [FREE Full text] [doi: 10.1186/s13058-019-1102-7] [Medline: 30736836]
4. Kay M, Santos J, Takane M. mHealth: New Horizons for Health Through Mobile Technologies. Global Observatory for eHealth Series - Volume 3. Geneva, Switzerland: World Health Organization; 2011. URL: https://apps.who.int/iris/bitstream/handle/10665/44607/9789241564250_eng.pdf?sequence=1 [accessed 2019-09-24]
5. Berkman LF, Glass T. Social integration, social networks, social support, and health. In: Berkman LF, Kawachi I, editors. Social Epidemiology. Oxford, UK: Oxford University Press; 2000:137-173.

6. Caulfield BM, Donnelly SC. What is Connected Health and why will it change your practice? QJM 2013 Aug;106(8):703-707. [doi: 10.1093/qjmed/hct114] [Medline: 23676416]
7. Kessel KA, Vogel MM, Kessel C, Bier H, Biedermann T, Friess H, et al. Mobile health in oncology: A patient survey about app-assisted cancer care. JMIR Mhealth Uhealth 2017 Jun 14;5(6):e81 [FREE Full text] [doi: 10.2196/mhealth.7689] [Medline: 28615159]
8. ICHOM. URL: https://www.ichom.org/ [accessed 2019-04-09] [WebCite Cache ID 77VUEETTe]
9. Common Terminology Criteria for Adverse Events (CTCAE): Version 5.0. Washington, DC: US Department of Health and Human Services; 2017 Nov 27. URL: https://ctep.cancer.gov/protocoldevelopment/electronic_applications/docs/ctcae_v5_quick_reference_5x7.pdf [accessed 2019-04-09] [WebCite Cache ID 77VUcrx7O]
10. Retzer A, Kyte D, Calman L, Glaser A, Stephens R, Calvert M. The importance of patient-reported outcomes in cancer studies. Expert Rev Qual Life Cancer Care 2018 May 09;3(2-3):65-71. [doi: 10.1080/23809000.2018.1472524]
11. Schuurhuizen CS, Braamse AM, Beekman AT, Bomhof-Roordink H, Bosmans JE, Cuijpers P, et al. Screening and treatment of psychological distress in patients with metastatic colorectal cancer: Study protocol of the TES trial. BMC Cancer 2015 Apr 17;15:302 [FREE Full text] [doi: 10.1186/s12885-015-1313-y] [Medline: 25903799]
12. Recklitis CJ, Syrjala KL. Provision of integrated psychosocial services for cancer survivors post-treatment. Lancet Oncol 2017 Jan;18(1):e39-e50. [doi: 10.1016/S1470-2045(16)30659-3]
13. Henson KE, Brock R, Charnock J, Wickramasinghe B, Will O, Pitman A. Risk of suicide after cancer diagnosis in England. JAMA Psychiatry 2019 Jan 01;76(1):51-60 [FREE Full text] [doi: 10.1001/jamapsychiatry.2018.3181] [Medline: 30476945]
14. Zaorsky NG, Zhang Y, Tuanquin L, Bluethmann SM, Park HS, Chinchilli VM. Suicide among cancer patients. Nat Commun 2019 Jan 14;10(1):207 [FREE Full text] [doi: 10.1038/s41467-018-08170-1] [Medline: 30643135]
15. Carver CS, Scheier MF, Weintraub JK. Assessing coping strategies: A theoretically based approach. J Pers Soc Psychol 1989;56(2):267-283. [doi: 10.1037/0022-3514.56.2.267]
16. Emery CF, Yang H, Frierson GM, Peterson LJ, Suh S. Determinants of physical activity among women treated for breast cancer in a 5-year longitudinal follow-up investigation. Psychooncology 2009 Apr;18(4):377-386. [doi: 10.1002/Pon.1519]
17. Osborn R, Demoncada A, Feuerstein M. Psychosocial interventions for depression, anxiety, and quality of life in cancer survivors: Meta-analyses. Int J Psychiatry Med 2006;36(1):13-34. [doi: 10.2190/EUFN-RV1K-Y3TR-FK0L] [Medline: 16927576]
18. Bower P, Gilbody S. Stepped care in psychological therapies: Access, effectiveness and efficiency. Narrative literature review. Br J Psychiatry 2005 Jan;186:11-17. [doi: 10.1192/bjp.186.1.11] [Medline: 15630118]
19. Leykin Y, Thekdi SM, Shumay DM, Muñoz RF, Riba M, Dunn LB. Internet interventions for improving psychological well-being in psycho-oncology: Review and recommendations. Psychooncology 2012 Sep;21(9):1016-1025. [doi: 10.1002/pon.1993] [Medline: 21608075]
20. Slev VN, Mistiaen P, Pasman HR, Verdonck-de Leeuw IM, van Uden-Kraan CF, Francke AL. Effects of eHealth for patients and informal caregivers confronted with cancer: A meta-review. Int J Med Inform 2016 Mar;87:54-67. [doi: 10.1016/j.ijmedinf.2015.12.013] [Medline: 26806712]
21. Escriva Boulley G, Leroy T, Bernetière C, Paquienseguy F, Desfriches-Doria O, Préau M. Digital health interventions to help living with cancer: A systematic review of participants' engagement and psychosocial effects. Psychooncology 2018 Dec;27(12):2677-2686. [doi: 10.1002/pon.4867] [Medline: 30152074]
22. McAlpine H, Joubert L, Martin-Sanchez F, Merolli M, Drummond KJ. A systematic review of types and efficacy of online interventions for cancer patients. Patient Educ Couns 2015 Mar;98(3):283-295. [doi: 10.1016/j.pec.2014.11.002] [Medline: 25535016]
23. Michie S, Yardley L, West R, Patrick K, Greaves F. Developing and evaluating digital interventions to promote behavior change in health and health care: Recommendations resulting from an international workshop. J Med Internet Res 2017 Jun 29;19(6):e232 [FREE Full text] [doi: 10.2196/jmir.7126] [Medline: 28663162]
24. Torous J, Andersson G, Bertagnoli A, Christensen H, Cuijpers P, Firth J, et al. Towards a consensus around standards for smartphone apps and digital mental health. World Psychiatry 2019 Feb;18(1):97-98 [FREE Full text] [doi: 10.1002/wps.20592] [Medline: 30600619]
25. Giunti G, Giunta DH, Guisado-Fernandez E, Bender JL, Fernandez-Luque L. A biopsy of breast cancer mobile applications: State of the practice review. Int J Med Inform 2018 Feb;110:1-9 [FREE Full text] [doi: 10.1016/j.ijmedinf.2017.10.022] [Medline: 29331247]
26. World Report on Disability. Geneva, Switzerland: World Health Organization; 2011. URL: http://whqlibdoc.who.int/publications/2011/9789240685215_eng.pdf?ua=1 [accessed 2019-09-24]
27. Cormie P, Zopf EM, Zhang X, Schmitz KH. The impact of exercise on cancer mortality, recurrence, and treatment-related adverse effects. Epidemiol Rev 2017 Jan 01;39(1):71-92. [doi: 10.1093/epirev/mxx007] [Medline: 28453622]
28. Kokila G, Smitha T. Cancer and physical activity. J Oral Maxillofac Pathol 2017;21(1):4-7 [FREE Full text] [doi: 10.4103/jomfp.JOMFP_20_17] [Medline: 28479678]
29. Dennett AM, Peiris CL, Shields N, Prendergast LA, Taylor NF. Cancer survivors awaiting rehabilitation rarely meet recommended physical activity levels. Rehabil Oncol 2018;36(4):214-222. [doi: 10.1097/01.REO.0000000000000132]

30. Rizzo A. The role of exercise and rehabilitation in the cancer care plan. J Adv Pract Oncol 2016 Apr;7(3):339-342 [FREE Full text] [Medline: 29152403]
31. 2018 Physical Activity Guidelines Advisory Committee. 2018 Physical Activity Guidelines Advisory Committee Scientific Report. Washington, DC: US Department of Health and Human Services; 2018. URL: https://health.gov/paguidelines/second-edition/report/pdf/PAG_Advisory_Committee_Report.pdf [accessed 2019-09-24]
32. Browall M, Mijwel S, Rundqvist H, Wengström Y. Physical activity during and after adjuvant treatment for breast cancer: An integrative review of women's experiences. Integr Cancer Ther 2018 Mar;17(1):16-30 [FREE Full text] [doi: 10.1177/1534735416683807] [Medline: 28008778]
33. Lowe SS, Danielson B, Beaumont C, Watanabe SM, Baracos VE, Courneya KS. Associations between objectively measured physical activity and quality of life in cancer patients with brain metastases. J Pain Symptom Manage 2014 Sep;48(3):322-332. [doi: 10.1016/j.jpainsymman.2013.10.012] [Medline: 24630754]
34. Clifford B, Mizrahi D, Sandler C, Barry B, Simar D, Wakefield C, et al. Barriers and facilitators of exercise experienced by cancer survivors: A mixed methods systematic review. Support Care Cancer 2018 Mar;26(3):685-700. [doi: 10.1007/s00520-017-3964-5] [Medline: 29185105]
35. Patel S, Park H, Bonato P, Chan L, Rodgers M. A review of wearable sensors and systems with application in rehabilitation. J Neuroeng Rehabil 2012 Apr 20;9(1):21 [FREE Full text] [doi: 10.1186/1743-0003-9-21] [Medline: 22520559]
36. Horak F, King L, Mancini M. Role of body-worn movement monitor technology for balance and gait rehabilitation. Phys Ther 2015 Mar 11;95(3):461-470 [FREE Full text] [doi: 10.2522/ptj.20140253] [Medline: 25504484]
37. Beg MS, Gupta A, Stewart T, Rethorst CD. Promise of wearable physical activity monitors in oncology practice. J Oncol Pract 2017 Feb;13(2):82-89 [FREE Full text] [doi: 10.1200/JOP.2016.016857] [Medline: 28387544]
38. Wang Q, Markopoulos P, Yu B, Chen W, Timmermans A. Interactive wearable systems for upper body rehabilitation: A systematic review. J Neuroeng Rehabil 2017 Mar 11;14(1):20 [FREE Full text] [doi: 10.1186/s12984-017-0229-y] [Medline: 28284228]
39. Treacy D, Hassett L, Schurr K, Chagpar S, Paul SS, Sherrington C. Validity of different activity monitors to count steps in an inpatient rehabilitation setting. Phys Ther 2017 May 01;97(5):581-588. [doi: 10.1093/ptj/pzx010] [Medline: 28339904]
40. Hultman E, Sjöholm H, Jäderholm-Ek I, Krynicki J. Evaluation of methods for electrical stimulation of human skeletal muscle in situ. Pflugers Arch 1983 Jul;398(2):139-141. [doi: 10.1007/bf00581062] [Medline: 6622220]
41. Maffiuletti NA. Physiological and methodological considerations for the use of neuromuscular electrical stimulation. Eur J Appl Physiol 2010 Sep 15;110(2):223-234. [doi: 10.1007/s00421-010-1502-y] [Medline: 20473619]
42. Banerjee P. Can electrical muscle stimulation of the legs produce cardiovascular exercise? J Clin Exp Cardiolog 2011;02(05):1-3. [doi: 10.4172/2155-9880.1000134]
43. O'Connor D, Caulfield B. The application of neuromuscular electrical stimulation (NMES) in cancer rehabilitation: Current prescription, pitfalls, and future directions. Support Care Cancer 2018 Nov;26(11):3661-3663. [doi: 10.1007/s00520-018-4269-z] [Medline: 29789948]
44. O'Connor D, Caulfield B, Lennon O. The efficacy and prescription of neuromuscular electrical stimulation (NMES) in adult cancer survivors: A systematic review and meta-analysis. Support Care Cancer 2018 Dec;26(12):3985-4000. [doi: 10.1007/s00520-018-4342-7] [Medline: 30022346]
45. O'Connor D, Fernandez MM, Signorelli G, Valero P, Caulfield B. Personalised and progressive neuromuscular electrical stimulation (NMES) in patients with cancer: A clinical case series. Support Care Cancer 2019 Oct 8;27(10):3823-3831. [doi: 10.1007/s00520-019-04679-3] [Medline: 30734089]
46. Giggins OM, Persson U, Caulfield B. Biofeedback in rehabilitation. J Neuroeng Rehabil 2013 Jun 18;10(1):60 [FREE Full text] [doi: 10.1186/1743-0003-10-60] [Medline: 23777436]
47. Chiviacowsky S, Wulf G. Feedback after good trials enhances learning. Res Q Exerc Sport 2007 Mar;78(2):40-47. [doi: 10.1080/02701367.2007.10599402] [Medline: 17479573]
48. Sigrist R, Rauter G, Riener R, Wolf P. Augmented visual, auditory, haptic, and multimodal feedback in motor learning: A review. Psychon Bull Rev 2013 Feb;20(1):21-53. [doi: 10.3758/s13423-012-0333-8] [Medline: 23132605]
49. Constantinescu G, Loewen I, King B, Brodt C, Hodgetts W, Rieger J. Designing a mobile health app for patients with dysphagia following head and neck cancer: A qualitative study. JMIR Rehabil Assist Technol 2017 Mar 24;4(1):e3 [FREE Full text] [doi: 10.2196/rehab.6319] [Medline: 28582245]
50. Cugelman B. Gamification: What it is and why it matters to digital health behavior change developers. JMIR Serious Games 2013 Dec 12;1(1):e3 [FREE Full text] [doi: 10.2196/games.3139] [Medline: 25658754]
51. Kato PM, Cole SW, Bradlyn AS, Pollock BH. A video game improves behavioral outcomes in adolescents and young adults with cancer: A randomized trial. Pediatrics 2008 Aug;122(2):e305-e317. [doi: 10.1542/peds.2007-3134] [Medline: 18676516]
52. Kappen D, Mirza-Babaei P, Nacke L. Gamification of older adults? Physical activity: An eight-week study. In: Proceedings of the 51st Hawaii International Conference on System Sciences. 2018 Presented at: 51st Hawaii International Conference on System Sciences; January 2-6, 2018; Waikoloa Village, HI p. 1207-1216. [doi: 10.24251/hicss.2018.149]

53. Litman L, Rosen Z, Spierer D, Weinberger-Litman S, Goldschein A, Robinson J. Mobile exercise apps and increased leisure time exercise activity: A moderated mediation analysis of the role of self-efficacy and barriers. J Med Internet Res 2015 Aug 14;17(8):e195 [FREE Full text] [doi: 10.2196/jmir.4142] [Medline: 26276227]
54. Bruggers C, Baranowski S, Beseris M, Leonard R, Long D, Schulte E, et al. A prototype exercise-empowerment mobile video game for children with cancer, and its usability assessment: Developing digital empowerment interventions for pediatric diseases. Front Pediatr 2018;6:69 [FREE Full text] [doi: 10.3389/fped.2018.00069] [Medline: 29686977]
55. Colón-Semenza C, Latham NK, Quintiliani LM, Ellis TD. Peer coaching through mHealth targeting physical activity in people with Parkinson disease: Feasibility study. JMIR Mhealth Uhealth 2018 Feb 15;6(2):e42 [FREE Full text] [doi: 10.2196/mhealth.8074] [Medline: 29449201]
56. Nasi G, Cucciniello M, Guerrazzi C. The performance of mHealth in cancer supportive care: A research agenda. J Med Internet Res 2015 Feb 13;17(1):e9 [FREE Full text] [doi: 10.2196/jmir.3764] [Medline: 25720295]
57. Ohri N, Kabarriti R, Bodner WR, Mehta KJ, Shankar V, Halmos B, et al. Continuous activity monitoring during concurrent chemoradiotherapy. Int J Radiat Oncol Biol Phys 2017 Apr 01;97(5):1061-1065 [FREE Full text] [doi: 10.1016/j.ijrobp.2016.12.030] [Medline: 28332990]
58. Vassbakk-Brovold K, Kersten C, Fegran L, Mjåland O, Mjåland S, Seiler S, et al. Cancer patients participating in a lifestyle intervention during chemotherapy greatly over-report their physical activity level: A validation study. BMC Sports Sci Med Rehabil 2016;8:10 [FREE Full text] [doi: 10.1186/s13102-016-0035-z] [Medline: 27099757]
59. Gartner. 2019. IT glossary: Internet of Things (IoT) URL: https://www.gartner.com/it-glossary/internet-of-things/ [accessed 2019-04-09] [WebCite Cache ID 77VLeX1Bm]
60. Zheng Y, Ding X, Poon CC, Lo BP, Zhang H, Zhou X, et al. Unobtrusive sensing and wearable devices for health informatics. IEEE Trans Biomed Eng 2014 May;61(5):1538-1554. [doi: 10.1109/TBME.2014.2309951] [Medline: 24759283]
61. Dickinson R, Hall S, Sinclair J, Bond C, Murchie P. Using technology to deliver cancer follow-up: A systematic review. BMC Cancer 2014 May 03;14:311 [FREE Full text] [doi: 10.1186/1471-2407-14-311] [Medline: 24885758]
62. Shen J, Naeim A. Telehealth in older adults with cancer in the United States: The emerging use of wearable sensors. J Geriatr Oncol 2017 Nov;8(6):437-442. [doi: 10.1016/j.jgo.2017.08.008] [Medline: 28888556]
63. Leaf Healthcare. Leaf patient monitoring system: Innovating pressure ulcer prevention URL: http://www.leafhealthcare.com/pdfs/LH_SO_Pressure-Ulcer-Prevention_1560AC_041817.pdf [accessed 2019-04-09] [WebCite Cache ID 77VVQQYUA]
64. Wearable Technology Insights. 2015 Apr 03. Bra helps to detect breast cancer URL: https://www.wearabletechnologyinsights.com/articles/7625/bra-helps-to-detect-breast-cancer [accessed 2019-04-09] [WebCite Cache ID 77VVgcUYW]
65. Yang Y, Gao W. Wearable and flexible electronics for continuous molecular monitoring. Chem Soc Rev 2019 Mar 18;48(6):1465-1491. [doi: 10.1039/C7CS00730B]
66. Constantinescu G, Jeong J, Li X, Scott D, Jang K, Chung H, et al. Epidermal electronics for electromyography: An application to swallowing therapy. Med Eng Phys 2016 Aug;38(8):807-812. [doi: 10.1016/j.medengphy.2016.04.023] [Medline: 27255865]
67. Tang Z, Sekine M, Tamura T, Yoshida M, Chen W. A chair for cuffless real-time estimation of systolic blood pressure based on pulse transit time. In: Proceedings of the 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). 2015 Presented at: 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC); August 25-29, 2015; Milan, Italy p. 5118-5121. [doi: 10.1109/EMBC.2015.7319543]
68. Lucchiari C, Masiero M, Pravettoni G. Methods for nutrition monitoring in cancer patients: A cognitive perspective. Ecancermedicalscience 2012;6:259 [FREE Full text] [doi: 10.3332/ecancer.2012.259] [Medline: 22837764]
69. Kogias K, Andreadis I, Dalakleidi K, Nikita KS. A two-level food classification system for people with diabetes mellitus using convolutional neural networks. In: Proceedings of the 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). 2018 Presented at: 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC); July 18-21, 2018; Honolulu, HI p. 2603-2606. [doi: 10.1109/EMBC.2018.8512839]
70. Shih PC, Han K, Poole ES, Rosson MB, Carroll JM. Use and adoption challenges of wearable activity trackers. In: Proceedings of the iConference 2015. 2015 Presented at: iConference 2015; March 24-27, 2015; Newport Beach, CA p. 1-12.
71. Epstein DA, Caraway M, Johnston C, Ping A, Fogarty J, Munson SA. Beyond Abandonment to Next Steps: Understanding and Designing for Life after Personal Informatics Tool Use. In: Proc SIGCHI Conf Hum Factor Comput Syst. 2016 May Presented at: SIGCHI Conference on Human Factors in Computing Systems; May 7-12, 2016; San Jose, CA p. 1109-1113 URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5428074/ [doi: 10.1145/2858036.2858045]
72. Lazar A, Koehler C, Tanenbaum J, Nguyen D. Why we use and abandon smart devices. In: Proceedings of the ACM International Joint Conference on Pervasive and Ubiquitous Computing. 2015 Presented at: ACM International Joint Conference on Pervasive and Ubiquitous Computing; September 7-11, 2015; Osaka, Japan p. 635-646. [doi: 10.1145/2750858.2804288]
73. Davis SW, Oakley-Girvan I. Achieving value in mobile health applications for cancer survivors. J Cancer Surviv 2017 Aug;11(4):498-504. [doi: 10.1007/s11764-017-0608-1] [Medline: 28342093]

74. Wesley KM, Fizur PJ. A review of mobile applications to help adolescent and young adult cancer patients. Adolesc Health Med Ther 2015;6:141-148 [FREE Full text] [doi: 10.2147/AHMT.S69209] [Medline: 26316835]
75. Klein M, Mogles N, van Wissen A. Intelligent mobile support for therapy adherence and behavior change. J Biomed Inform 2014 Oct;51:137-151 [FREE Full text] [doi: 10.1016/j.jbi.2014.05.005] [Medline: 24858491]
76. Hawkins RP, Kreuter M, Resnicow K, Fishbein M, Dijkstra A. Understanding tailoring in communicating about health. Health Educ Res 2008 Jun;23(3):454-466 [FREE Full text] [doi: 10.1093/her/cyn004] [Medline: 18349033]
77. Puszkiewicz P, Roberts AL, Smith L, Wardle J, Fisher A. Assessment of cancer survivors' experiences of using a publicly available physical activity mobile application. JMIR Cancer 2016 May 31;2(1):e7 [FREE Full text] [doi: 10.2196/cancer.5380] [Medline: 28410168]
78. Jacobs M, Johnson J, Mynatt ED. MyPath: Investigating breast cancer patients' use of personalized health information. Proc ACM Hum Comput Interact 2018 Nov 01;2(CSCW):1-21. [doi: 10.1145/3274347]
79. Robertson MC, Tsai E, Lyons EJ, Srinivasan S, Swartz MC, Baum ML, et al. Mobile health physical activity intervention preferences in cancer survivors: A qualitative study. JMIR Mhealth Uhealth 2017 Jan 24;5(1):e3 [FREE Full text] [doi: 10.2196/mhealth.6970] [Medline: 28119278]
80. Roberts AL, Potts HW, Koutoukidis DA, Smith L, Fisher A. Breast, prostate, and colorectal cancer survivors' experiences of using publicly available physical activity mobile apps: Qualitative study. JMIR Mhealth Uhealth 2019 Jan 04;7(1):e10918 [FREE Full text] [doi: 10.2196/10918] [Medline: 30609982]
81. Brier MJ, Schwartz LA, Kazak AE. Psychosocial, health-promotion, and neurocognitive interventions for survivors of childhood cancer: A systematic review. Health Psychol 2015 Feb;34(2):130-148. [doi: 10.1037/hea0000119] [Medline: 25133839]
82. Enwald HP, Huotari ML. Preventing the obesity epidemic by second generation tailored health communication: An interdisciplinary review. J Med Internet Res 2010 Jun 28;12(2):e24 [FREE Full text] [doi: 10.2196/jmir.1409] [Medline: 20584698]
83. Kennedy CM, Powell J, Payne TH, Ainsworth J, Boyd A, Buchan I. Active assistance technology for health-related behavior change: An interdisciplinary review. J Med Internet Res 2012 Jun 14;14(3):e80 [FREE Full text] [doi: 10.2196/jmir.1893] [Medline: 22698679]
84. Lustria ML, Cortese J, Noar SM, Glueckauf RL. Computer-tailored health interventions delivered over the Web: Review and analysis of key components. Patient Educ Couns 2009 Feb;74(2):156-173. [doi: 10.1016/j.pec.2008.08.023] [Medline: 18947966]
85. op den Akker H, Jones VM, Hermens HJ. Tailoring real-time physical activity coaching systems: A literature survey and model. User Model User-adapt Interact 2014 Jun 25;24(5):351-392. [doi: 10.1007/s11257-014-9146-y]
86. Kuijpers W, Groen WG, Aaronson NK, van Harten WH. A systematic review of Web-based interventions for patient empowerment and physical activity in chronic diseases: Relevance for cancer survivors. J Med Internet Res 2013 Feb 20;15(2):e37 [FREE Full text] [doi: 10.2196/jmir.2281] [Medline: 23425685]
87. Groen W, Kuijpers W, Oldenburg H, Wouters M, Aaronson N, van Harten WH. Empowerment of cancer survivors through information technology: An integrative review. J Med Internet Res 2015 Nov 27;17(11):e270 [FREE Full text] [doi: 10.2196/jmir.4818] [Medline: 26614438]
88. Kenfield S, Van Blarigan EL, Ameli N, Lavaki E, Cedars B, Paciorek A, et al. Feasibility, acceptability, and behavioral outcomes from a technology-enhanced behavioral change intervention (Prostate 8): A pilot randomized controlled trial in men with prostate cancer. Eur Urol 2019 Jun;75(6):950-958. [doi: 10.1016/j.eururo.2018.12.040] [Medline: 30638635]
89. Krist AH, Woolf SH. A vision for patient-centered health information systems. JAMA 2011 Jan 19;305(3):300-301 [FREE Full text] [doi: 10.1001/jama.2010.2011] [Medline: 21245186]
90. Kourou K, Exarchos TP, Exarchos KP, Karamouzis MV, Fotiadis DI. Machine learning applications in cancer prognosis and prediction. Comput Struct Biotechnol J 2015;13:8-17 [FREE Full text] [doi: 10.1016/j.csbj.2014.11.005] [Medline: 25750696]
91. Huang C, Clayton EA, Matyunina LV, McDonald LD, Benigno BB, Vannberg F, et al. Machine learning predicts individual cancer patient responses to therapeutic drugs with high accuracy. Sci Rep 2018 Nov 06;8(1):16444 [FREE Full text] [doi: 10.1038/s41598-018-34753-5] [Medline: 30401894]
92. Lee J. American Society of Clinical Oncology. 2018 Nov 13. Artificial-intelligence smartphone app significantly reduced severity of cancer patients' pain and hospital admissions URL: https://www.asco.org/about-asco/press-center/news-releases/artificial-intelligence-smartphone-app-significantly-reduced [accessed 2019-04-09] [WebCite Cache ID 77VW6G4Ww]
93. De Silva D, Ranasinghe W, Bandaragoda T, Adikari A, Mills N, Iddamalgoda L, et al. Machine learning to support social media empowered patients in cancer care and cancer treatment decisions. PLoS One 2018;13(10):e0205855 [FREE Full text] [doi: 10.1371/journal.pone.0205855] [Medline: 30335805]
94. Research Centre for Data Analytics and Cognition. GitHub. Patient Reported Information Multidimensional Exploration (PRIME) URL: https://github.com/tharindurb/PRIME [accessed 2019-04-09] [WebCite Cache ID 77VWR75YC]
95. Jung H, Chung K. Knowledge-based dietary nutrition recommendation for obese management. Inf Technol Manage 2015 Mar 14;17(1):29-42. [doi: 10.1007/s10799-015-0218-4]

96. Hors-Fraile S, Benjumea F, Hernández L, Ruiz F, Fernandez-Luque L. arXiv. 2016. Design of two combined health recommender systems for tailoring messages in a smoking cessation app URL: https://arxiv.org/ftp/arxiv/papers/1608/1608.07192.pdf [accessed 2018-11-28] [WebCite Cache ID 74FVpjHgQ]
97. Blobel B, Stegwee R. Standards and solutions for eHealth interoperability. Eur J Biomed Inform (Praha) 2012;8(3):1-2. [doi: 10.24105/ejbi.2012.08.3.1]
98. European Commission. 2017 Jan 25. €1 Million for more interoperability in eHealth URL: https://ec.europa.eu/digital-single-market/en/news/eu1-million-more-interoperability-ehealth [accessed 2019-04-09] [WebCite Cache ID 77VY2hKS0]
99. Dogac A. Interoperability in eHealth systems. In: Proceedings of the 38th International Conference on Very Large Databases. 2012 Presented at: 38th International Conference on Very Large Databases; August 27-31, 2012; Istanbul, Turkey p. 2026-2027 URL: http://vldb.org/pvldb/vol5/p2026_asumandogac_vldb2012.pdf [doi: 10.14778/2367502.2367568]
100. HL7 International. FHIR overview URL: https://www.hl7.org/fhir/overview.html [accessed 2019-04-09] [WebCite Cache ID 77VswR0oF]
101. Bender D, Sartipi K. HL7 FHIR: An agile and RESTful approach to healthcare information exchange. In: Proceedings of the 26th IEEE International Symposium on Computer-Based Medical Systems. 2013 Presented at: 26th IEEE International Symposium on Computer-Based Medical Systems; June 20-22, 2013; Porto, Portugal URL: https://ieeexplore.ieee.org/abstract/document/6627810 [doi: 10.1109/CBMS.2013.6627810]
102. Gøeg KR, Rasmussen RK, Jensen L, Wollesen CM, Larsen S, Pape-Haugaard LB. A future-proof architecture for telemedicine using loose-coupled modules and HL7 FHIR. Comput Methods Programs Biomed 2018 Jul;160:95-101. [doi: 10.1016/j.cmpb.2018.03.010] [Medline: 29728251]
103. Hong J, Morris P, Seo J. Interconnected personal health record ecosystem using IoT cloud platform and HL7 FHIR. In: Proceedings of the 2017 IEEE International Conference on Healthcare Informatics (ICHI). 2017 Presented at: 2017 IEEE International Conference on Healthcare Informatics (ICHI); August 23-26, 2017; Park City, UT p. 362-367 URL: https://conferences.computer.org/ichi/2017/papers/4881a362.pdf [doi: 10.1109/ichi.2017.82]
104. Bennett AV, Dueck AC, Mitchell SA, Mendoza TR, Reeve BB, Atkinson TM, National Cancer Institute PRO-CTCAE Study Group. Mode equivalence and acceptability of tablet computer-, interactive voice response system-, and paper-based administration of the US National Cancer Institute's Patient-Reported Outcomes version of the Common Terminology Criteria for Adverse Events (PRO-CTCAE). Health Qual Life Outcomes 2016 Feb 19;14:24 [FREE Full text] [doi: 10.1186/s12955-016-0426-6] [Medline: 26892667]
105. Basch E, Deal AM, Kris MG, Scher HI, Hudis CA, Sabbatini P, et al. Symptom monitoring with patient-reported outcomes during routine cancer treatment: A randomized controlled trial. J Clin Oncol 2016 Feb 20;34(6):557-565 [FREE Full text] [doi: 10.1200/JCO.2015.63.0830] [Medline: 26644527]
106. Basch E, Deal AM, Dueck AC, Scher HI, Kris MG, Hudis C, et al. Overall survival results of a trial assessing patient-reported outcomes for symptom monitoring during routine cancer treatment. JAMA 2017 Jul 11;318(2):197-198 [FREE Full text] [doi: 10.1001/jama.2017.7156] [Medline: 28586821]
107. ICT Data and Statistics Division, Telecommunication Development Bureau of ITU. Measuring the Information Society. Geneva, Switzerland: International Telecommunication Union (ITU); 2012. URL: https://www.itu.int/en/ITU-D/Statistics/Documents/publications/mis2012/MIS2012_without_Annex_4.pdf [accessed 2019-09-24]
108. Greenspun H, Coughlin S. mHealth in an mWorld: How Mobile Technology is Transforming Health Care. Washington, DC: Deloitte Center for Health Solutions; 2012. URL: https://www2.deloitte.com/content/dam/Deloitte/us/Documents/life-sciences-health-care/us-lhsc-mhealth-in-an-mworld-103014.pdf [accessed 2019-04-09] [WebCite Cache ID 77VcCCkds]
109. Jahns R, Houck P. Mobile Health Market Report 2013-2017: The Commercialization of mHealth Applications (Vol 3). Berlin, Germany: Research 2 Guidance; 2013 Mar 04. URL: https://research2guidance.com/wp-content/uploads/2015/08/Mobile-Health-Market-Report-2013-2017-Preview.pdf [accessed 2019-09-24]
110. Topol EJ. High-performance medicine: The convergence of human and artificial intelligence. Nat Med 2019 Jan 7;25(1):44-56. [doi: 10.1038/s41591-018-0300-7] [Medline: 30617339]
111. Council of the European Union, European Parliament. Official Journal of the European Communities. 1995 Nov 23. Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data URL: https://publications.europa.eu/en/publication-detail/-/publication/775a4724-2086-4a06-9213-1a4e6489053b/language-en [accessed 2019-09-24]
112. Kvedar JC, Colman C, Cella G. The privacy trade-offs. In: The Internet of Healthy Things. Boston, MA: Partners Connected Health; 2015:170-180.
113. Council of the European Union, European Parliament. Official Journal of the European Union. 2016 Apr 27. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) URL: https://publications.europa.eu/en/publication-detail/-/publication/3e485e15-11bd-11e6-ba9a-01aa75ed71a1/language-en [accessed 2019-09-24]

Abbreviations

AI: artificial intelligence
CATCH: Cancer: Activating Technology for Connected Health
C-CDA: Consolidated Clinical Document Architecture
COST: European Cooperation in Science and Technology
CTCAE: Common Terminology Criteria for Adverse Events
EHR: electronic health record
ENJECT: European Network for the Joint Evaluation of Connected Health Technologies
EURO-CAS: eHealth Interoperability Conformity Assessment Scheme for Europe
FCT: Fundação para a Ciência e a Tecnologia
FHIR: Fast Healthcare Interoperability Resources
HL7: Health Level Seven
ICHOM: International Consortium for Health Outcomes Measurement
IHTSDO: International Health Terminology Standards Development Organisation
mHealth: mobile health
NMES: neuromuscular electrical stimulation
OASIS: Organization for the Advancement of Structured Information Standards
OMG: Object Management Group
PRIME: Patient-Reported Information Multidimensional Exploration
PROM: patient-reported outcome measure
QoL: quality of life
RESTful: representational state transfer
SMS: short message service

Edited by M Tilney; submitted 12.04.19; peer-reviewed by T Kovacevic; comments to author 10.07.19; revised version received 07.08.19; accepted 25.08.19; published 29.10.19.

Please cite as:
Signorelli GR, Lehocki F, Mora Fernández M, O'Neill G, O'Connor D, Brennan L, Monteiro-Guerra F, Rivero-Rodriguez A, Hors-Fraile S, Munoz-Penas J, Bonjorn Dalmau M, Mota J, Oliveira RB, Mrinakova B, Putekova S, Muro N, Zambrana F, Garcia-Gomez JM
A Research Roadmap: Connected Health as an Enabler of Cancer Patient Support
J Med Internet Res 2019;21(10):e14360
URL: http://www.jmir.org/2019/10/e14360/
doi:10.2196/14360
PMID:31663861

Original Paper

The Effect of Robot Attentional Behaviors on User Perceptions and Behaviors in a Simulated Health Care Interaction: Randomized Controlled Trial

Deborah L Johanson[1,2], MHealthPsych; Ho Seok Ahn[2,3], PhD; Bruce A MacDonald[2,3], PhD; Byeong Kyu Ahn[2,3], MScEng; JongYoon Lim[2,3], BComSc; Euijun Hwang[2,3], BComSc; Craig J Sutherland[2,3], PhD; Elizabeth Broadbent[1,2], PhD

[1]Department of Psychological Medicine, Faculty of Medical and Health Sciences, The University of Auckland, Auckland, New Zealand

[2]Centre for Automation and Robotic Engineering Science, The University of Auckland, Auckland, New Zealand

[3]Department of Electrical, Computer, and Software Engineering, Faculty of Engineering, The University of Auckland, Auckland, New Zealand

Corresponding Author:
Elizabeth Broadbent, PhD
Department of Psychological Medicine
Faculty of Medical and Health Sciences
The University of Auckland
Level 12 Auckland Hospital, Room 12005
2 Park Road
Auckland, 1023
New Zealand
Phone: 64 99236756
Email: e.broadbent@auckland.ac.nz

Related Article:

This is a corrected version. See correction statement: https://www.jmir.org/2020/3/e18362/

Abstract

Background: For robots to be effectively used in health applications, they need to display appropriate social behaviors. A fundamental requirement in all social interactions is the ability to engage, maintain, and demonstrate attention. Attentional behaviors include leaning forward, self-disclosure, and changes in voice pitch.

Objective: This study aimed to examine the effect of robot attentional behaviors on user perceptions and behaviors in a simulated health care interaction.

Methods: A parallel randomized controlled trial with a 1:1:1:1 allocation ratio was conducted. We randomized participants to 1 of 4 experimental conditions before engaging in a scripted face-to-face interaction with a fully automated medical receptionist robot. Experimental conditions included a self-disclosure condition, voice pitch change condition, forward lean condition, and neutral condition. Participants completed paper-based postinteraction measures relating to engagement, perceived robot attention, and perceived robot empathy. We video recorded interactions and coded for participant attentional behaviors.

Results: A total of 181 participants were recruited from the University of Auckland. Participants who interacted with the robot in the forward lean and self-disclosure conditions found the robot to be significantly more stimulating than those who interacted with the robot in the voice pitch or neutral conditions (P=.03). Participants in the forward lean, self-disclosure, and neutral conditions found the robot to be significantly more interesting than those in the voice pitch condition (P<.001). Participants in the forward lean and self-disclosure conditions spent significantly more time looking at the robot than participants in the neutral condition (P<.001). Significantly, more participants in the self-disclosure condition laughed during the interaction (P=.01), whereas significantly more participants in the forward lean condition leant toward the robot during the interaction (P<.001).

Conclusions: The use of self-disclosure and forward lean by a health care robot can increase human engagement and attentional behaviors. Voice pitch changes did not increase attention or engagement. The small effects with regard to participant perceptions are potentially because of the limitations in self-report measures or a lack of comparison for most participants who had never

interacted with a robot before. Further research could explore the use of self-disclosure and forward lean using a within-subjects design and in real health care settings.

(J Med Internet Res 2019;21(10):e13667) doi:10.2196/13667

KEYWORDS

robotics; health care robotics; social interaction; engagement; social intelligence

Introduction

Background

The use of social robots in home and health care environments is fast becoming a reality [1,2]. Although much robotics research is focused on the technical capabilities of robots, it is also important that research considers the behaviors of robots to ensure that interactions between humans and robots are successful. Consideration of robot social behaviors is perhaps even more salient when considering health care robots, which may be interacting with potentially vulnerable individuals on a daily basis. For interactions between patients and health care robots to be successful, these robots will need to behave in a way that is not only useful but also acceptable and comfortable [3]. One way to inform research investigating appropriate robot social behaviors in human-robot interactions is to consider the social behaviors that lead to successful human interactions.

Attentional behaviors are an important group of human social behaviors, fundamental to ensuring successful interactions. Attentional behaviors include the ability to not only demonstrate attention but also to engage and maintain the attention of others. As put by Zhao et al [4], "mutual attentiveness leads to an experience of connectedness" (p. 515). Given the importance of attention in human social interactions, it is critical that researchers investigating the social aspects of health care robots explore human attentional research to inform potential research within this area. A number of researchers have, in fact, taken this approach, researching several key human attentional behaviors within the context of human-robot interactions. One of these key behaviors is eye gaze.

Eye gaze is crucial to establish human joint attention, which, in turn, is a critical aspect of human learning, communication, and social interaction [5,6]. In a health care context, the appropriate use of eye gaze by a physician has been found to be associated with increased patient satisfaction and increased patient ratings of physician empathy, physician attention, and physician warmth [7,8]. In human-robot interactions, robot eye gaze has been found to increase human attention and engagement and facilitate comprehension of robot communication [9-11].

Despite growing research into the importance of robot attentional behaviors, a number of key human attentional behaviors are yet to be explored within the context of human-robot interaction, especially in health care. Three such attentional behaviors include the use of self-disclosure, voice pitch changes, and a forward lean. These attentional behaviors have been found to be important in human social interactions and in interactions between patients and health care professionals. The following section of this paper describes previous research examining the use of self-disclosure, voice pitch, and forward lean in human interactions. In the instance that research has been done exploring one of these behaviors in the context of human-robot interactions, this is presented.

Self-Disclosure

Self-disclosure refers to the *act of revealing personal information about oneself to another* and is recognized as central to the process of building close relationships [12]. Research into the use of self-disclosure in human interactions has found that self-disclosure is more effective when negatively skewed. For example, Zhao et al [4] found that in conversations between 2 individuals, previously unknown to each other, self-disclosure was often in the form of personally negative statements (eg, "I'm always late for the bus"), and that these statements were then often met with similarly negative statements ("Me too!"). The authors state that this seemingly superficial conversation tool increased mutual gaze among participants and often lead to more intimate conversation. Failure to reciprocate negative self-disclosure can lead to decreases in feelings of rapport [13].

The study of robot self-disclosure and its effect on human-robot interactions is limited. In the research that has been undertaken so far, the use of self-disclosure by a robot has been shown to increase users' ratings of a robot's agency and experience [12], stabilize users' anxiety about a robot's communication capacity [14], increase users' perceptions of a robot's likability, and decrease users' feelings of control [15]. No research to date has examined the effect of robot self-disclosure on human engagement or attention. Patient attention and engagement are critical to patient satisfaction and adherence; therefore, it is important that further research is conducted to investigate whether self-disclosure from a robot can influence attention to the robot in a health care application [16].

Forward Body Lean

It is equally important to examine nonverbal attentional behaviors as it is to examine verbal attentional behaviors. Eye gaze is one example of an important nonverbal attentional behavior, used by a listener to engage and demonstrate attention toward a speaker.

The use of a forward body lean by a listener is another salient way to display attention, interest, and agreeance toward a speaker [17-19]. Leaning forward toward another individual to display attention is an almost automatic behavior and is even found to be used in those communicating through sign language [19].

Given the importance of using a forward body lean to demonstrate attention in human interactions, it is perhaps surprising to note that there has been no research undertaken examining the effect of robot forward body lean, in the context

of human-robot interactions, on any outcome variables. There has been some research on robot forward neck tilt, which, when used alongside expressive facial movements, has been shown to aid human recognition of robot emotions [20], human comprehension of robot behaviors [21], and facilitate turn-taking [21]. The robot used in these studies, however, was not a health care robot and was made up of a *head* and *neck* with no *body*. The lack of research on forward body lean by any robot, particularly a health care robot, represents an important gap in our knowledge.

Voice Pitch Changes

A person's voice pitch, or in other words, how low or high a voice is in frequency [22], conveys a range of information to others, such as gender and emotional affect [22,23]. Voice pitch has even been found to influence perceptions of attractiveness, with research finding that men rate woman with high-pitched voices as more attractive than those with low-pitched voices [24]. Research in the area of verbal communication has found that individuals can determine the personality traits of others with considerable accuracy, purely through patterns of speech, such as speed and voice pitch [25].

Voice pitch and voice pitch changes are an important part of attending behaviors and essential for communication [26]. Voice pitch changes allow a speaker to place emphasis on certain words, infuse emotion into specific phrases, and influence comprehension through the use of inflection (eg, in the case of a statement or question) to initiate and sustain the attention of others [26,27]. Therefore, key is the use of voice pitch changes in attention and communication, for example, individuals often exaggerate voice pitch changes when storytelling to hold the attention of their audience [28]. Other research investigating voice pitch in human interactions provides further support for the effect of voice pitch on attention. A recent study found that retention of content in long-term memory was higher when individuals listened to voices using high and low voice pitches, as opposed to a medium voice pitch [29]. This finding was independent of whether individuals listened to natural voices or voices that had been manipulated. These studies provide a rationale for examining the effects of both high and low robot voice pitch changes on human-robot interactions.

Although previous research in human-robot interactions has investigated robot voice–related variables, such as robot voice gender [30], robot voice age [31], and robot voice human likeness [32], in regard to user outcomes, only 1 study to date has explored the effect of robot voice pitch in the context of human-robot interactions. This study by Niculescu et al [22] compared a robot with a high voice pitch against a robot with a low voice pitch, finding that the robot with the high voice pitch was rated by participants as significantly more likable and attractive, with a better *personality*. In addition, the interaction with the robot with the high-pitched voice was rated more exciting, entertaining, and enjoyable.

Justification for Research in Health Care

A model of robot-patient interaction proposes that behaviours that are important in physician-patient communication may also be important in communication between healthcare robots and patients [33]. Self-disclosure, forward lean, and voice pitch are key behaviors in good physician-patient communication, as detailed below. Therefore, these behaviors are likely to be important in interactions between health care robots and patients to establish a good rapport. However, research is needed to test this hypothesis.

Physician-Patient Communication Theory

Doctor-patient or physician-patient communication theory posits that the way in which a physician communicates with a patient is just as important as the information they provide for patient outcomes. Physician-patient communication is key to building rapport with patients and central to the delivery of appropriate health care [33,34]. Effective physician verbal and nonverbal communication has been found to decrease patient anxiety and psychological distress [35,36], facilitate patient understanding of medical information [34], and increase patient satisfaction [33].

The central goals of effective physician-patient communication are to facilitate the establishment of rapport, enable exchange of health-related information, and promote patient involvement in treatments plans and health-related decision making [33,34,36,37]. To accomplish these goals, a physician needs to act in a way that demonstrates their attention to the patient. In fact, the Toronto Consensus Agreement recommends that all physicians should actively attend to patients to encourage full expression of health concerns, without interruption or premature closure of conversation [37].

Forward Lean in Health Care

As previously discussed, use of a forward body lean is 1 way to demonstrate attention and establish rapport. In a health care context, forward lean is used by clinicians to demonstrate attention or *active listening* to patients [29]. Indeed, a study by Sharpley and Sagris [38] found that the use of forward lean by a counsellor was associated with increased client ratings of rapport. In research examining physician-patient interactions, the use of forward lean by a physician was found to be associated with positive patient perceptions of physician empathy, respect, and genuineness [39]. A systematic review by Beck et al [40] found physician forward lean to be among the behaviors that were significantly associated with increases in physician-patient rapport, increased patient satisfaction, and increased patient understanding. Owing to the fact that forward lean is used as a way of demonstrating attention across many cultures, its use is recommended to health care professionals as appropriate for use with most patients [29].

Self-Disclosure in Health Care

Although many physicians are trained not to self-disclose, a recent systematic review found self-disclosure was routinely used by physicians in clinical practice [41]. This same review found that, when used appropriately, self-disclosure by a physician had the potential to significantly increase patient satisfaction and physician-patient rapport. In clinicians, appropriate use of self-disclosure usually involves the disclosure of personal information to a patient that is relevant to the therapeutic process [42]. In this way, self-disclosure is able to

demonstrate to a patient that they have been heard, as well as creating a feeling of shared experience.

Voice Pitch in Health Care

In clinicians, voice pitch (or voice tone) is used as a way to demonstrate attention and empathy to patients [29]. In a review of doctor-patient communication [43], researchers found that patients were less satisfied with their consultation when their physician used a negative voice tone or had tension in their tone. In a later study of surgeon's malpractice history, surgeons who were perceived to have a dominant voice tone (ie, deep and loud) were more likely to have been sued by patients compared with surgeons with a less dominant voice pitch (ie, higher and softer) [44].

This Study

The human medical receptionist is the first point of contact for patients seeking medical assistance in specialist and general practice clinics [45]. As such, the communication behaviors of receptionists can significantly impact patients [46,47]. It is therefore important that researchers investigate which aspects of human behaviors are appropriate for medical receptionist robots to ensure interactions are successful.

As detailed above, a number of human attentional behaviors have yet to be fully explored within the context of health care human-robot interactions. In regard to self-disclosure, a handful of studies have investigated the effect of robot self-disclosure on user outcomes such as anxiety [15], perceived robot likability [14], and perceived robot mind attribution [12]. However, no studies have investigated the effect of robot self-disclosure on any user attentional outcomes, such as engagement (attention), perceived robot attention, and perceived robot empathy, or user attentional behaviors in health care. In regard to forward body lean, no study could be found investigating the effect of robot forward body lean on user attentional behaviors or any other outcome.

Although 1 study has compared the effect of a robot using a high voice pitch against the effect of a robot using a low voice pitch, no study could be found investigating the effect of a robot using both high and low voice pitch changes within a single human-robot interaction. Given the importance of self-disclosure, forward lean, and voice pitch changes in both human social interactions and interactions between health care professionals and patients, it is critical that robotics researchers investigate these attentional behaviors in interactions between robots and human users. Research in this area could potentially inform the future design and implementation of not only social robots but also, more specifically, health care robots.

Study Objectives and Hypotheses

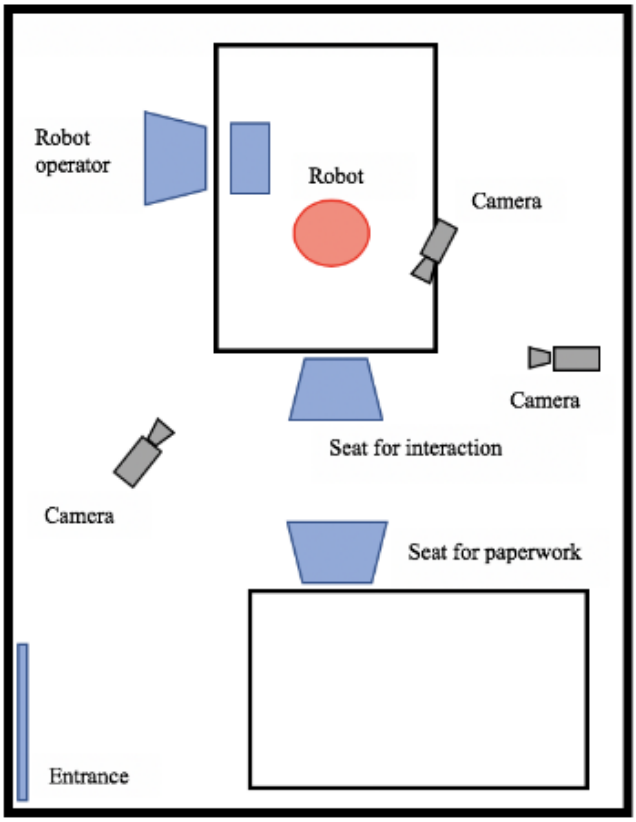
The aim of this research was to investigate the effects of robot self-disclosure, forward lean, and voice pitch changes on user perceptions and attentional behaviors in a health care context. We hypothesized that, compared with a neutral condition, these robot behaviors would increase participants' perceptions of engagement, robot empathy, and robot attention and increase participants' own attentional behaviors.

Methods

Experimental Design

A between-subjects experimental study was conducted at the University of Auckland, Auckland, New Zealand. Participants completed baseline measures before being randomized to 1 of the 4 experimental groups (ie, self-disclosure, voice pitch, forward lean, or neutral condition). Following the interaction with the robot, postinteraction measures were completed. The interaction between each participant and the robot was video recorded from 3 different angles to allow for coding of participant behaviors (Figure 1).

Figure 1. Experimental setup.



Participants

Participants were recruited through flyers posted at the University of Auckland's campuses and through emails to students. Eligibility criteria were being 16 years of age or older and fluent in English. Written informed consent was obtained.

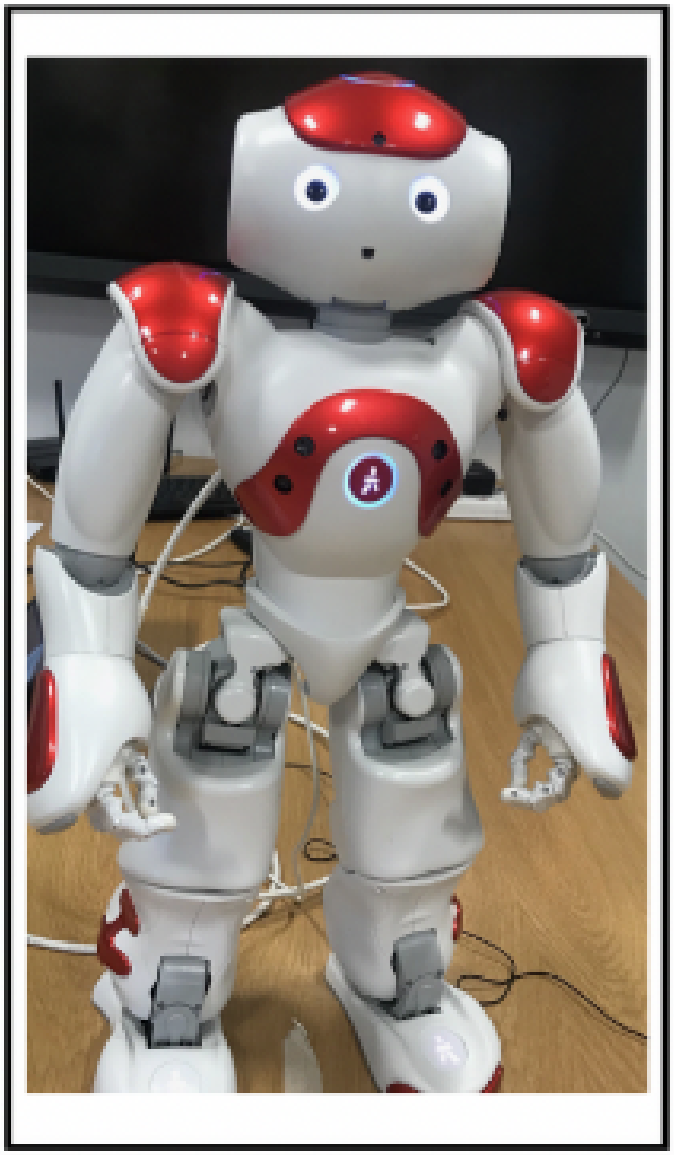
The Robot

A Nao robot (Softbank, Japan) was chosen for this study, as it was able to meet the requirements of the experimental conditions in regard to forward body lean, spoken conversation, and voice pitch changes. The Nao robot is an autonomous, programmable, humanoid robot, able to perform a variety of physical movements and speech patterns (Figure 2). A single Nao robot was used, with each participant interacting with the robot on an individual basis.

Procedures

Once randomized to 1 of the 4 experimental conditions, participants were asked to imagine that they were attending their current general practitioner's office, and the Nao robot was the robot receptionist. The participant was provided with a script for use during the interaction with the robot. This script was identical for all conditions and instructed participants to undertake a variety of tasks during their interaction with the Nao robot (Multimedia Appendix 1). The tasks undertaken during the interaction included activities that account for over 90% of the interactions between human medical receptionists and patients, such as checking in for a doctor's appointment and collecting a prescription [47]. Once the robot operator selected the appropriate experimental interaction, the robot introduced itself as "Nao the Receptionist Robot" before enquiring as to how it could be of help to the participant. The Nao robot used identical speech responses and prompts across all conditions, with the exception of the self-disclosure condition.

Figure 2. Nao: the medical receptionist robot.



In the self-disclosure condition, the Nao robot stated, "I'm a little nervous about this task" after introducing itself to the participant. In addition, when a participant advised that they could not remember the name of the doctor they were seeing (as per the Participant Scenario Information Sheet, Multimedia Appendix 1), the robot stated, "no problem, I forget things too sometimes" before continuing the scripted interaction. In the forward body lean condition, the Nao robot leaned (approximately 20°) forward toward a participant when he or she was speaking, maintaining a neutral standing position during the rest of the interaction (Figure 3).

In the voice pitch condition, the Nao robot both *increased and decreased* its voice pitch within the single interaction. The Nao robot *decreased* its voice pitch (by 15%) when apologizing and advising a participant that their script was not available for collection. This lowering of robot voice pitch was intended to display robot *sadness* to the participant in relation to being unable to assist with their request. Studies investigating vocal emotion recognition have shown that sadness is best recognized through voice alone when a lower voice pitch is used [48,49] (compared with a neutral). The robot *increased* its voice pitch (by 15%) when advising a participant, it was "no problem" in regard to helping the participant with the name of the doctor they were seeing and when stating "I hope you have a nice day" at the end of the interaction. This increase in robot voice pitch was intended to display robot *happiness* to the participant in relation to being able to assist the participant and in wishing them a good day. Studies investigating vocal emotion recognition have shown that happiness is best recognized through voice alone when a higher voice pitch is used [48,49] (compared with a neutral). Finally, in the neutral condition, the Nao robot maintained a neutral standing position and neutral voice pitch throughout the interaction, with no self-disclosure.

Measures

Participant Engagement

A Likert scale was developed using an adaption of the *stimulation* items from the McGill Friendship Questionnaire [50], along with an adaption of the engagement items used in the human-robot engagement study by Snider et al [51] (Multimedia Appendix 2). The Cronbach alpha for the combination of these Likert items was found to be .86, showing

excellent reliability. Therefore, the scores from these items were added to create a total *participant engagement* score.

In addition, pair-choice items were developed using an adaption of the *stimulation* (paired) items from the AttrakDiff user experience tool created by Hassenzahl et al [52]. All pair-choice items were analyzed separately.

Adaptions of both the McGill Friendship Questionnaire and the AttrakDiff user experience tool have been used previously in human-robot interaction research [22,53].

Figure 3. The Nao robot—from a neutral standing to forward lean position.

Perceived Robot Attention

No scale was found measuring human perceptions of robot attention or attentiveness. Therefore, a new measure was created using an adaption of the *stimulation* items from the McGill Friendship Questionnaire [50], along with an adaption of the engagement items used in the human-robot interaction study by Snider et al [51] (Multimedia Appendix 2). The Cronbach alpha for this measure was .89, showing excellent reliability. Thus, the scores from all items were added to create a total *perceived robot attention* score.

Perceived Robot Empathy

An adaption of the McGill Friendship Questionnaire [50] was used along with an adaption of the consultation and relational empathy measure [54] (Multimedia Appendix 2). The consultation and relational empathy measure assesses patient perceived empathy in relation to clinical encounters and has been found to be both valid and reliable across clinical settings [55,56]. An adaption of the McGill Friendship Questionnaire was used in research investigating perceived robot empathy by Leite et al [53]. The Cronbach alpha for this combined measure was found to be .82. After removal of the item, "I think Nao had fun during this interaction," Cronbach alpha increased to .89. Therefore, the scores from all remaining items were added to create a total *perceived robot empathy* score.

Observer Ratings

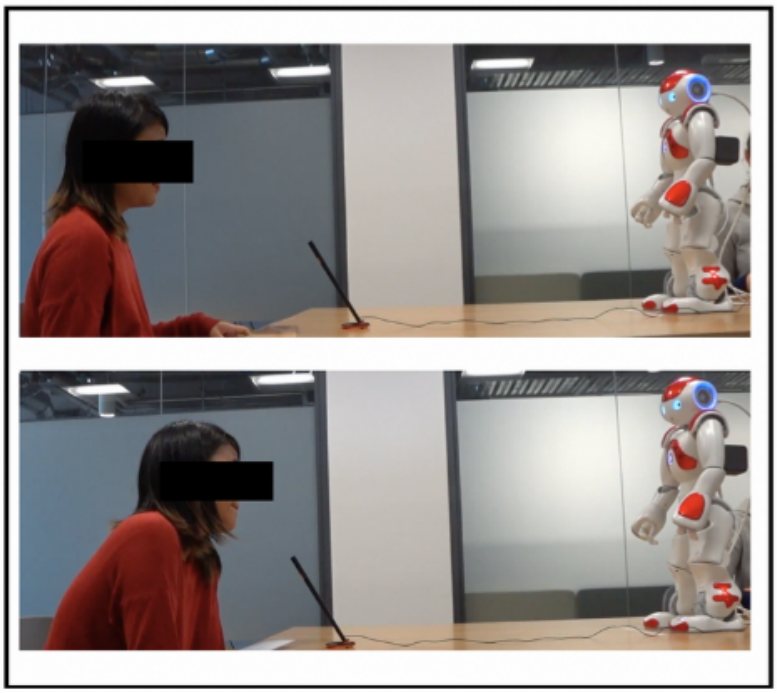
The video ratings were used in addition to self-reports to measure participant attention and engagement. Eye gaze and forward lean were used to measure attention, and smiling and laughter were used to measure engagement.

Video recordings were viewed to determine the overall time (in seconds) that each participant spent looking at the Nao robot during the interaction. Using this time and the total interaction time, a percentage was able to be determined in regard to the time spent looking at the Nao robot. Each video recording was coded in regard to whether or not a participant leant toward the Nao robot during the course of the interaction. A *forward lean* was identified by placing a visual marker on the participants' midback (during the video playback) when they had settled into the experimental chair and then watching to see if the participant leaned toward the robot, forward from the marker, at any time

during the interaction (Figure 4). As many participants sat down and immediately leant forward (resting their forearms on the table), this posture was not identified as a forward lean unless active forward leaning past this *neutral* point was identified. In addition, many participants leant forward into the microphone (that was positioned on the desk) when they spoke, and therefore, this forward leaning movement was disregarded.

Each of the video recordings was then viewed to determine the overall time (in seconds) that each participant spent smiling during the interaction with the Nao robot. As with time spent looking at the robot, this time was used to give a percentage of time spent smiling at the robot for each participant. Each video recording was also coded in regard to whether or not a participant laughed during the course of the interaction with the Nao robot. The coding in regard to whether or not a participant *laughed* during the interaction with Nao was based on the laughter intensity scale developed by Law et al [57].

Figure 4. Participant moving from neutral to forward lean position.



Statistical Analyses

Power

The sample size was determined by a power analysis using the G*Power program created by Faul et al [58]. The following parameters were selected: an alpha error probability of .05, power of 0.80, and an effect size (f) of 0.31. The effect size used was based on an average of the effect sizes found in previous studies examining robot social behaviors in the context of human-robot interactions [14,22]. The analysis revealed a total sample size of 180 participants (45 participants per experimental condition) would be required.

Interaction Times

A 1-way analysis of variance (ANOVA) was used to analyze the total interaction times to determine if time spent with the Nao robot differed significantly between conditions. This analysis was undertaken as a significant difference in interaction times between groups would represent a potential confound.

Self-Report Measures

One-way ANOVAs were used to analyze the total participant engagement and total perceived robot empathy scores, with a posthoc (Tukey) test used to compare the conditions pairwise when an ANOVA was found to be significant. Fisher exact tests were used to analyze the results of the pair-choice engagement items. A Kruskal-Wallis test was performed to analyze total perceived robot attention because of the data being found to violate normality.

Video Analyses

One-way ANOVAs were used to analyze the percentage time participants spent looking at the robot (participant eye gaze) as well as the percentage time participants spent smiling during the interaction with the robot, with a posthoc (Tukey) test used to compare the conditions individually when an ANOVA was found to be significant. Fisher exact tests were used to determine if any significant differences existed between conditions in regard to whether or not participants laughed during the interaction with the robot, as well as whether or not participants leaned toward the robot during the interaction.

All analyses were performed using the SPSS, version 22.

Results

Manipulation Check

A manipulation check was performed using a convenience sample (n=10). Participants undertaking the manipulation check were asked to view 4 separate video recordings of the robot (1 video of each condition) and completed a brief measure after each video. The measure used asked participants to indicate

whether the robot in the video demonstrated a forward lean, used self-disclosure statements, used voice pitch changes, or *none of the above*. Before viewing the recordings, participants were verbally given the following definition to identify robot self-disclosure: "self-disclosure is the act of revealing personal information about oneself to another [12], please indicate if you feel the robot has revealed personal information about itself in any of the following recordings."

Furthermore, 100% (10/10) of participants were able to accurately identify self-disclosure and forward lean behaviors in the robot. Of 10 participants, 1 (10%) confused the voice pitch and neutral conditions, but the remaining 9 participants (90%) were able to accurately identify voice pitch changes and neutral behaviors in the robot.

Participants

In total, 181 participants took part in this study. Participants were predominantly female (112/181, 61.9%) and ranged in age from 17 to 80 years (mean 25.8, SD 10.21). Most participants were students (n=139), followed by part-time employees (n=20), full-time employees (n=19), and those who were retired or currently unemployed (n=4). Participants mainly identified as being of New Zealand European ethnicity (n=57), followed by Chinese (n=37), Indian (n=29), Korean (n=5), Maori (n=4), Samoan (n=3), and Tongan (n=1). In addition, 49 participants identified as having an ethnicity other than those listed on the baseline questionnaire form. Most participants (148/181, 81.8%,) advised that they had never before interacted with any kind of robot.

Interaction Times

Interactions ran from 145 to 284 seconds in total, with a mean total interaction time of 182 seconds. There were no significant differences found between the means of the neutral (187.98), forward lean (178.26), self-disclosure (180.72), and voice pitch (180.67) conditions in regard to total interaction time with the Nao robot ($F_{3,173}$=1.14; P=.34).

Participant Perceived Engagement Scores

A 1-way ANOVA of total participant engagement scores ($F_{3,177}$=1.420; P=.24) found no significant difference between the means of the neutral (mean 26.96, SD 5.80), forward lean (mean 27.38, SD 5.87), self-disclosure (mean 28.22, SD 5.39), and voice pitch (mean 25.84, SD 5.26) conditions.

A Fisher exact test of pair-choice engagement items found that participants in the voice pitch condition were significantly more likely to rate Nao as boring (as opposed to interesting) compared with the neutral, self-disclosure, and forward lean groups (χ^2_3=10.3; P<.001; n=179). A large effect size (Cramer's V=.29) was found for this item. No significant differences were found between the conditions in regard to participant rating of *unimaginative versus creative* (χ^2_3=2.4; P=.54; n=178), *cautious versus bold* (χ^2_3=5.6; P=.13; n=172), *innovative versus conservative* (χ^2_3=0.3; P=.97; n=180), *dull versus absorbing* (χ^2_3=5.5; P=.14; n=177), or *novel versus conservative* (χ^2_3=3.4; P=.34; n=178). Participants in the voice pitch and neutral conditions were significantly more likely to rate Nao as unstimulating (as opposed to stimulating) compared with the self-disclosure and forward lean groups (χ^2_3=8.8; P=.03; n=176). A medium effect size (Cramer's V=.22) was found for this item.

Perceived Robot Empathy

A 1-way ANOVA of total perceived robot empathy scores ($F_{3,175}$=1.89; P=.13) found no significant differences between the means of the forward lean (mean 44.23, SD 6.72), self-disclosure (mean 43.83, SD 7.32), voice pitch (mean 41.33, SD 6.92), and neutral (mean 41.95, SD 6.38) conditions.

Perceived Robot Attention

A Kruskal-Wallis test of total perceived robot attention scores (χ^2_3=1.1; P=.78; n=181) found no significant difference between the mean rank (MR) scores of the forward lean (MR=94.44), self-disclosure (MR=94.63), voice pitch (MR=90.02), and neutral (MR=84.82) conditions.

Participant Behaviors

In total, 174 video recordings were coded for analysis. Of the 7 participant interactions that were not analyzed, 5 were excluded because of the technical difficulties with recording equipment, and 2 were excluded because of participants' refusal to be recorded during the interaction.

Participant Eye Gaze

There was a significant difference in the percentage time participants spent looking at the Nao robot during the interaction ($F_{3,173}$=8.13; P<.001). Participants in the forward lean (mean 78.80, SD 8.98) condition spent significantly more time looking at the robot compared with participants in the neutral (mean 69.14, SD 10.96) and voice pitch (mean 73.30, SD 9.88) conditions. Participants in the self-disclosure (mean 76.30, SD 8.78) condition were found to have spent significantly more time looking at the Nao robot during the interaction compared with participants in the neutral condition. A medium to large effect size (η2=.13) was found for this condition. All other comparisons were nonsignificant.

Participant Use of Forward Lean

There was a significant difference between the conditions in regard to whether or not participants leaned toward the Nao robot during the interaction (χ^2_3=22.1; P<.001; n=174). Significantly, more participants leaned toward the Nao robot in the forward lean condition, 67% (31/46), compared with the self-disclosure, 47% (20/42), voice pitch, 39% (17/43), and neutral, 18% (8/43) conditions. A large effect size (Cramer's V=.36) was found for this condition.

Participant Smiling

There were no significant differences in the percentage of time participants spent smiling during the interaction with the robot ($F_{3,173}$=0.801; P=.50). The means of the neutral (mean 9.35, SD 9.28), forward lean (mean 9.20, SD 8.25), self-disclosure (mean 11.98, SD 9.73), and voice pitch (mean 10.70, SD 10.90) conditions did not significantly differ in regard to the percentage of time participants spent smiling during the interaction.

Participant Laughing

There was a significant difference between groups in whether participants laughed or not (χ^2_3=12.0; P=.01; n=174). Significantly, more participants laughed in the self-disclosure condition, 47% (20/42), compared with the forward lean, 21% (10/46), voice pitch, 20% (9/43), and neutral, 18% (8/43) conditions. A medium effect size (Cramer's V=.26) was found for this condition.

Discussion

Principal Findings

The forward body lean and self-disclosure robot behaviors showed significant effects on both self-reported outcomes and observed behaviors compared with the voice pitch and neutral conditions. Participants in the robot forward body lean condition spent significantly more time looking at the robot compared with participants in the voice pitch and neutral conditions. They were also more likely to lean forward toward the robot than those who interacted with a neutral robot and reported the robot was significantly more stimulating than participants in the voice pitch and neutral conditions. Participants in the self-disclosure condition also spent significantly more time looking at the robot compared with participants in the neutral condition, and significantly, more participants in the self-disclosure condition laughed during the interaction compared with participants in the forward lean, voice pitch, and neutral conditions. The voice pitch condition had no effects or even slightly negative effects compared with the other conditions with participants in the forward lean, self-disclosure, and neutral conditions finding the robot to be significantly more interesting than participants in the voice condition.

There were no significant differences found between groups in regard to self-reported engagement, perceived robot empathy, or perceived robot attention. A potential explanation for these results may be found in the use of a between-subjects study design. This design may have resulted in a lack of comparison for the majority of participants (148/181) who had never interacted with any kind of robot before this experiment. Another potential explanation is that the novelty or excitement of a *first encounter* with a robot may have created some ceiling effects in regard to the measures used. Certainly, many participants expressed excitement in regard to interacting with the Nao robot, and the total scores for participant perceived robot attention and participant perceived engagement were positively skewed regardless of condition. Finally, some of the null outcomes may be because of the use of self-report measures that rely on the memory of the interaction.

In contrast, the study did find differences between groups in behavioral measures of attention and engagement. Differences between groups in eye gaze, laughing, and forward lean suggest that participant engagement was significantly higher in the forward lean and self-disclosure groups compared with the voice pitch and neutral conditions. Behavioral measures offer some advantages over self-reports, as they are less prone to memory and social desirability bias and can be more sensitive, valid, and reliable [59].

Comparison With Prior Work

Forward Lean

This study is novel in many regards. It is the first study to investigate the use of robot forward body lean in a human-robot interaction and the first to show that robot forward body lean can positively influence users' attentional behaviors and self-reported stimulation. This finding is salient in terms of social robotics research, as it represents a simple robot nonverbal behavior that may be used to increase user engagement. Increasing engagement between health care robots and users is fundamental to ensuring positive interactions and important in establishing user attention and comprehension. Previous literature examining interactions between health care professionals and patients demonstrates the importance of a forward lean by a clinician to demonstrate *active listening* [29]. By leaning forward toward participants when it spoke, the Nao robot may have been perceived by participants as actively listening to their questions and responses. It may be therefore that the increased eye gaze and forward body lean behaviors observed in participants in the forward body lean condition represent a form of reciprocated attention, or mirroring, toward the robot. This theory is supported by the fact that participants in the forward lean condition found the interaction with the Nao robot to be significantly more stimulating and significantly more interesting when compared with the neutral condition. As the videos in this study were not analyzed in regard to *when* forward leaning occurred, just whether or not a forward leaning behavior was observed, reciprocated attention, or mirroring, could not be ascertained.

In contrast to previous research showing an association between physician forward body lean and patient perceptions of physician empathy [39], the robot's use of forward body lean did not result in increased user perceptions of robot empathy. As discussed above, this finding may be because of the use of a between-subjects design, resulting in a lack of comparison for the majority of participants who had never interacted with a robot before. A within-subjects design in which participants interact with both the neutral robot and forward lean robot may have shown different outcomes in regard to user perceived robot empathy.

Self-Disclosure

Although previous research in robotics has investigated the effect of robot self-disclosure in terms of decreasing user anxiety, increasing user perceived robot likability, and increasing user perceived robot mind attribution [12,14,15], this is the first study to show that self-disclosure by a robot can increase user attentional and engagement behaviors. Although participant *self-reported* measures reflecting *perceived* engagement did not differ significantly when comparing the self-disclosure condition with the neutral condition, participants in the self-disclosure condition did spend significantly more time *looking* at the robot during the interaction compared with participants in the neutral condition. This observed increase in robot-directed eye gaze behaviors indicated that there was an increased level of participant attention and engagement with the Nao robot in the self-disclosure condition.

Once again, this increase in user engagement gives important insights into the basic social behaviors that a social or health care robot can display to increase engagement and attention during interactions. As discussed earlier, increased robot-directed eye gaze behaviors, such as seen in this experiment, have also been found in research examining the effect of self-disclosure in human social interactions [4].

Participant laughing also indicates a higher level of engagement during the self-disclosure condition compared with the neutral condition. Participants in the self-disclosure condition were found to laugh significantly more than participants in any of the other conditions. This was an incidental finding and not one we had expected to make when beginning this experiment. Participants in this condition were most likely to laugh when the Nao robot stated, "Don't worry, I forget things too sometimes" in response to a participant stating that they had forgotten the name of their doctor. It may be that, as well as acting as self-disclosure statement, participants perceived this statement as being humorous. The use of humor in health care environments is gaining attention as a useful therapeutic tool for the facilitation of positive patient health outcomes [60]. This in turn has generated a fledgling area of social robotics research investigating the effect of robot humor on human-robot interactions [61,62]. Given the significant response by participants (in the form of laughter) to the Nao robots *forgetfulness*, it appears that robot humor is an area worth exploring.

Voice Pitch

Although previous research has examined the effect of robot voice gender, robot voice age, and robot voice human likeness [30-32] in regard to user outcomes, only 1 other study has investigated robot voice pitch within the context of a human-robot interaction [22]. In this study [22], researchers used a within-subjects design to compare participant interactions between a robot using a *high* voice pitch against interactions with a robot using a *low* voice pitch, finding that the robot with the high voice pitch was rated more positively by participants. The robot in the voice pitch condition of our study was programmed to use both high *and* low voice pitch changes within a single interaction. The lack of significant results in our study in regard to the voice pitch condition may suggest that voice pitch changes are beneficial only if they are in the higher range. Another potential explanation for these findings may be inadequate frequency and/or distinction in regard to the voice pitch changes used by the robot. Indeed, the manipulation check did show that 1 of 10 participants had difficulty distinguishing the voice pitch condition from the neutral condition.

Contribution to Existing Literature

This study demonstrates the importance of robot forward lean and self-disclosure in increasing human attentional and engagement behaviors in a health care application. The ability of a robot to attract and sustain human attention is particularity important in health care environments. Patients in these settings need to pay attention to advice, reminders, and recommendations of health care robots. The implications of this research are that social robot designers should consider the implementation of robot forward lean and self-disclosure to increase user attention, particularly with regard to health care robots. These results extend the literature demonstrating the importance of forward lean and self-disclosure in interactions between health professionals and patients to robots [38-41]. Furthermore, this research supports a robot-patient interaction model that proposes the importance of verbal and nonverbal behaviors to user outcomes [16].

Limitations

Similar to many studies in human-robot interaction, the participants were mainly students and relatively young. Younger people may be more positive and open in regard to interacting with a robot than older people, and therefore, the results may have limited generalizability to an older population. In addition, the study was conducted in a laboratory setting, and further research is needed in a more realistic setting. Another limitation is that the study used a scripted interaction. The behaviors and self-reported measures observed may have differed from what would have been observed had a natural conversation taken place between participants and the Nao robot. Finally, there is research to suggest that an individual's personality type may influence perceptions and attitudes toward robots [63]. As we did not use any measures of human personality type in this study, it is unknown whether participants' personality influenced outcomes. Nevertheless, randomization to groups should ensure that personality did not systematically differ between groups. A strength of the research is that participants were blinded to group allocation.

Future Work

Future research could consider the use of a within-subjects design to provide participants with a basis of comparison, allowing for greater insight into the effects of specific robot attentional behaviors within an experimental context. Furthermore, the use of natural speech, rather than a scripted interaction, may allow participants to concentrate fully on the interaction with the robot, as opposed to splitting attention between the robot and script.

Given that the real-life implementation of a robot medical receptionist would certainly take place within a natural environment (such as a medical clinic), research exploring the use of a health care robot's attentional behaviors within the context of such an environment (eg, an actual doctor's clinic) is necessary. This research would not only provide potentially significant insight into the effect of robot attentional behaviors in naturalistic settings but may also allow for recruiting of a more mixed-age sample.

Finally, research investigating the effect of robot voice pitch changes in human-robot interactions may need to focus on using more distinct and/or frequent voice pitch changes to see significant effects. Future work could accentuate a health care robot's voice pitch changes and explore the effects on user perceptions of robot acceptability.

Conclusions

This study explored the effects of robot self-disclosure, robot forward lean, and robot voice pitch changes on user perceptions of engagement, robot attention, and robot empathy, as well as

user attentional behaviors. Robot self-disclosure and forward body lean resulted in significantly better self-reported outcomes and observed behaviors compared with the neutral condition. Robot voice pitch changes did not have positive effects, but more research is needed to further investigate this. Exploring the effect of human social behaviors, such as attentional behaviors, within the context of human-robot interactions in health care, represents a salient opportunity to inform future robotic design and implementation.

Acknowledgments

This work was supported by the Technology Innovation Program (10077553, Development of Social Robot Intelligence for Social Human-Robot Interaction of Service Robots) funded By the Ministry of Trade, Industry & Energy (MI, Korea). The authors would also like to acknowledge the University of Auckland Robotics Department and the CARES team for their ongoing support during this study.

Conflicts of Interest

None declared.

Editorial notice: This randomized study was not prospectively registered. The editor granted an exception of ICMJE rules for prospective registration of randomized trials because the risk of bias appears low and the study was considered formative. However, readers are advised to carefully assess the validity of any potential explicit or implicit claims related to primary outcomes or effectiveness.

Multimedia Appendix 1
Participant scenario information sheet.
[PDF File (Adobe PDF File), 254 KB - jmir_v21i10e13667_app1.pdf]

Multimedia Appendix 2
Postinteraction measures.
[PDF File (Adobe PDF File), 226 KB - jmir_v21i10e13667_app2.pdf]

Multimedia Appendix 3
CONSORT‐EHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 561 KB - jmir_v21i10e13667_app3.pdf]

References

1. Bemelmans R, Gelderblom GJ, Jonker P, de Witte L. Effectiveness of robot Paro in intramural psychogeriatric care: a multicenter quasi-experimental study. J Am Med Dir Assoc 2015 Nov 1;16(11):946-950. [doi: 10.1016/j.jamda.2015.05.007] [Medline: 26115817]
2. Robinson H, Macdonald B, Kerse N, Broadbent E. The psychosocial effects of a companion robot: a randomized controlled trial. J Am Med Dir Assoc 2013 Sep;14(9):661-667. [doi: 10.1016/j.jamda.2013.02.007] [Medline: 23545466]
3. Walters ML, Syrdal DS, Dautenhahn K, te Boekhorst R, Koay KL. Avoiding the uncanny valley: robot appearance, personality and consistency of behavior in an attention-seeking home scenario for a robot companion. Auton Robots 2007 Nov 20;24(2):159-178. [doi: 10.1007/s10514-007-9058-3]
4. Zhao R, Papangelis A, Cassell J. Towards a Dyadic Computational Model of Rapport Management for Human-Virtual Agent Interaction. In: Proceedings of the International Conference on Intelligent Virtual Agents. 2014 Presented at: IVA'14; August 26-29, 2014; Boston, MA p. 514-527. [doi: 10.1007/978-3-319-09767-1_62]
5. Mundy P, Newell L. Attention, joint attention, and social cognition. Curr Dir Psychol Sci 2007 Oct 1;16(5):269-274 [FREE Full text] [doi: 10.1111/j.1467-8721.2007.00518.x] [Medline: 19343102]
6. Rohlfing KJ, Longo MR, Bertenthal BI. Dynamic pointing triggers shifts of visual attention in young infants. Dev Sci 2012 May;15(3):426-435. [doi: 10.1111/j.1467-7687.2012.01139.x] [Medline: 22490182]
7. Griffith CH, Wilson JF, Langer S, Haist SA. House staff nonverbal communication skills and standardized patient satisfaction. J Gen Intern Med 2003 Mar;18(3):170-174 [FREE Full text] [doi: 10.1046/j.1525-1497.2003.10506.x] [Medline: 12648247]
8. Harrigan JA, Oxman TE, Rosenthal R. Rapport expressed through nonverbal behavior. J Nonverbal Behav 1985;9(2):95-110. [doi: 10.1007/BF00987141]
9. Karreman DE, Bradford GU, van Dijk EM, Lohse M, Evers V. Picking Favorites: The Influence of Robot Eye-Gaze on Interactions With Multiple Users. In: Proceedings of the International Conference on Intelligent Robots and Systems. 2013 Presented at: RSJ'13; November 3-7, 2013; Tokyo, Japan p. 3-7 URL: https://doi.org/10.1109/IROS.2013.6696342 [doi: 10.1109/IROS.2013.6696342]

10. Kirchner N, Alempijevic A, Dissanayake G. Nonverbal Robot-Group Interaction Using an Imitated Gaze Cue. In: Proceedings of the 6th International Conference on Human-Robot Interaction. 2011 Presented at: HRI'11; March 6-9, 2011; Lausanne, Switzerland p. 497-504 URL: https://doi.org/10.1145/1957656.1957824 [doi: 10.1145/1957656.1957824]
11. Moon A, Troniak DM, Gleeson B, Pan MK, Zheng M, Blumer BA, et al. Meet Me Where I'm Gazing: How Shared Attention Gaze Affects Human-Robot Handover Timing. In: Proceedings of the 2014 ACM/IEEE International Conference on Human-Robot Interaction. 2014 Presented at: HRI'14; March 3-6, 2014; Bielefeld, Germany p. 334-341 URL: https://doi.org/10.1145/2559636.2559656 [doi: 10.1145/2559636.2559656]
12. Eyssel F, Wullenkord R, Nitsch V. The Role of Self-Disclosure in Human-Robot Interaction. In: Proceedings of the International Symposium on Robot and Human Interactive Communication. 2017 Presented at: RO-MAN'17; August 28-September 1, 2017; Lisbon, Portugal. [doi: 10.1109/ROMAN.2017.8172413]
13. Bronstein I, Nelson N, Livnat Z, Ben-Ari R. Rapport in negotiation: the contribution of the verbal channel. J Confl Resolut 2012 Jun 14;56(6):1089-1115. [doi: 10.1177/0022002712448913]
14. Nomura T, Kawakami K. Relationships Between Robot's Self-Disclosures and Human's Anxiety Toward Robots. In: Proceedings of the International Conferences on Web Intelligence and Intelligent Agent Technology. 2011 Presented at: WIC'11; August 22-27, 2011; Lyon, France p. 22-27 URL: httphttps://ieeexplore.ieee.org/document/6040807 [doi: 10.1109/WI-IAT.2011.17]
15. Siino RM, Chung J, Hinds PJ. Colleague vs Tool: Effects of Disclosure in Human-Robot Collaboration. In: Proceedings of the 17th IEEE International Symposium on Robot and Human Interactive Communication. 2008 Presented at: RO-MAN'08; August 1-3, 2008; Munich, Germany URL: https://ieeexplore.ieee.org/document/4600725 [doi: 10.1109/ROMAN.2008.4600725]
16. Broadbent E, Johanson D, Shah J. A New Model to Enhance Robot-Patient Communication: Applying Insights from the Medical World. In: Proceedings of the International Conference on Social Robotics. 2018 Presented at: ICSR'18; November 28-30, 2018; Qingdao, China p. 308-317 URL: https://link.springer.com/chapter/10.1007/978-3-030-05204-1_30 [doi: 10.1007/978-3-030-05204-1_30]
17. Whitlock J, Lloyd-Richardson E. Psychology Today. 2018. Adolescent Communication 101: Tips for the Holidays URL: https://www.psychologytoday.com/nz/blog/youth-and-consequences/201812/adolescent-communication-101-tips-the-holidays
18. van der Kooij E, Crasborn O, Emmerik W. Explaining prosodic body leans in sign language of the Netherlands: pragmatics required. J Pragmat 2006 Oct;38(10):1598-1614. [doi: 10.1016/j.pragma.2005.07.006]
19. Wilbur RB, Patschke CG. Body leans and the marking of contrast in American sign language. J Pragmat 1998 Sep;30(3):275-303. [doi: 10.1016/S0378-2166(98)00003-4]
20. Breazeal C. Emotion and social humanoid robots. Int J Hum-Comput St 2003;59(1-2):119-155. [doi: 10.1016/S1071-5819(3)00018-1]
21. Breazeal C. Toward sociable robots. Robot Auton Syst 2003 Mar;42(3-4):167-175. [doi: 10.1016/S0921-8890(02)00373-1]
22. Niculescu A, van Dijk B, Nijholt A, Li H, See SL. Making social robots more attractive: the effects of voice pitch, humor and empathy. Int J Soc Robot 2013 Apr;5(2):171-191. [doi: 10.1007/s12369-012-0171-x]
23. Saxton TK, Mackey LL, McCarty K, Neave N. A lover or a fighter? Opposing sexual selection pressures on men's vocal pitch and facial hair. Behav Ecol 2016;27(2):512-519 [FREE Full text] [doi: 10.1093/beheco/arv178] [Medline: 27004013]
24. Collins SA, Missing C. Vocal and visual attractiveness are related in women. Anim Behav 2003 May;65(5):997-1004. [doi: 10.1006/anbe.2003.2123]
25. Kwon S, Choeh JY, Lee JW. User-personality classification based on the non-verbal cues from spoken conversations. Int J Comput Int Sys 2013 Aug;6(4):739-749. [doi: 10.1080/18756891.2013.804143]
26. Sommers-Flanagan R, Sommers-Flanagan J. Clinical Interviewing. Fourth Edition. Hoboken, NJ: Wiley; 2012.
27. Kort NS, Nagarajan SS, Houde JF. A bilateral cortical network responds to pitch perturbations in speech feedback. Neuroimage 2014 Feb 1;86:525-535 [FREE Full text] [doi: 10.1016/j.neuroimage.2013.09.042] [Medline: 24076223]
28. Theune M, Meijs K, Heylen D, Ordelman R. Generating expressive speech for storytelling applications. IEEE Trans Audio Speech Language Process 2006 Jul;14(4):1137-1144. [doi: 10.1109/TASL.2006.876129]
29. Helfrich H, Weidenbecher P. Impact of voice pitch on text memory. Swiss J Psychol 2011 Jan;70(2):85-93. [doi: 10.1024/1421-0185/a000042]
30. Crowelly CR, Villanoy M, Scheutzz M, Schermerhornz P. Gendered Voice and Robot Entities: Perceptions and Reactions of Male and Female Subjects. In: Proceedings of the International Conference on Intelligent Robots and Systems. 2009 Presented at: RSJ'09; October 10-15, 2009; St Louis, MO, USA URL: https://ieeexplore.ieee.org/document/5354204 [doi: 10.1109/IROS.2009.5354204]
31. Edwards C, Edwards A, Stoll B, Lin X, Massey N. Evaluations of an artificial intelligence instructor's voice: social identity theory in human-robot interactions. Comput Hum Behav 2019 Jan;90:357-362. [doi: 10.1016/j.chb.2018.08.027]
32. Trovato G, Ramos JG, Azevedo H, Moroni A, Magossi S, Ishii H, et al. Designing a Receptionist Robot: Effect of Voice and Appearance on Anthropomorphism. In: Proceedings of the International Symposium on Robot and Human Interactive Communication. 2015 Presented at: RO-MAN'15; August 31-September 4, 2015; Kobe, Japan. [doi: 10.1109/ROMAN.2015.7333573]

33. Ha JF, Longnecker N. Doctor-patient communication: a review. Ochsner J 2010;10(1):38-43 [FREE Full text] [doi: 10.3329/jbcps.v32i2.26036] [Medline: 21603354]
34. Travaline JM, Ruchinskas R, D'Alonzo GE. Patient-physician communication: why and how. J Am Osteopath Assoc 2005 Jan;105(1):13-18. [Medline: 15710660]
35. Butow PN, Kazemi JN, Beeney LJ, Griffin AM, Dunn SM, Tattersall MH. When the diagnosis is cancer: patient communication experiences and preferences. Cancer 1996 Jun 15;77(12):2630-2637. [doi: 10.1002/(SICI)1097-0142(19960615)77:12<2630::AID-CNCR29>3.0.CO;2-S] [Medline: 8640715]
36. Stewart M, Brown JB, Boon H, Galajda J, Meredith L, Sangster M. Evidence on patient-doctor communication. Cancer Prev Control 1999 Feb;3(1):25-30. [Medline: 10474749]
37. Simpson M, Buckman R, Stewart M, Maguire P, Lipkin M, Novack D, et al. Doctor-patient communication: the Toronto consensus statement. Br Med J 1991 Nov 30;303(6814):1385-1387 [FREE Full text] [doi: 10.1136/bmj.303.6814.1385] [Medline: 1760608]
38. Sharpley CF, Sagris A. When does counsellor forward lean influence client-perceived rapport? Brit J Guid Couns 2007 Oct 16;23(3):387-394 [FREE Full text] [doi: 10.1080/03069889508253696]
39. Hall JA, Harrigan JA, Rosenthal R. Nonverbal behavior in clinician—patient interaction. Appl Prev Psychol 1995 Dec;4(1):21-37. [doi: 10.1016/S0962-1849(05)80049-6]
40. Beck RS, Daughtridge R, Sloane PD. Physician-patient communication in the primary care office: a systematic review. J Am Board Fam Pract 2002;15(1):25-38 [FREE Full text] [Medline: 11841136]
41. Allen EF, Arroll B. Physician self-disclosure in primary care: a mixed methods study of GPs' attitudes, skills, and behaviour. Br J Gen Pract 2015 Sep;65(638):e601-e608 [FREE Full text] [doi: 10.3399/bjgp15X686521] [Medline: 26324497]
42. Audet CT, Everall RD. Therapist self-disclosure and the therapeutic relationship: a phenomenological study from the client perspective. Br J Guid Counc 2010 Aug;38(3):327-342. [doi: 10.1080/03069885.2010.482450]
43. Williams S, Weinman J, Dale J. Doctor-patient communication and patient satisfaction: a review. Fam Pract 1998 Oct;15(5):480-492. [doi: 10.1093/fampra/15.5.480] [Medline: 9848436]
44. Ambady N, Laplante D, Nguyen T, Rosenthal R, Chaumeton N, Levinson W. Surgeons' tone of voice: a clue to malpractice history. Surgery 2002 Jul;132(1):5-9. [doi: 10.1067/msy.2002.124733] [Medline: 12110787]
45. Neuwelt PM, Kearns RA, Browne AJ. The place of receptionists in access to primary care: challenges in the space between community and consultation. Soc Sci Med 2015 May;133:287-295. [doi: 10.1016/j.socscimed.2014.10.010] [Medline: 25455478]
46. Ward J, McMurray R. The unspoken work of general practitioner receptionists: a re-examination of emotion management in primary care. Soc Sci Med 2011 May;72(10):1583-1587. [doi: 10.1016/j.socscimed.2011.03.019] [Medline: 21501912]
47. Hewitt H, McCloughan L, McKinstry B. Front desk talk: discourse analysis of receptionist-patient interaction. Br J Gen Pract 2009 Aug;59(565):e260-e266 [FREE Full text] [doi: 10.3399/bjgp09X453774] [Medline: 22751251]
48. Gold R, Butler P, Revheim N, Leitman DI, Hansen JA, Gur RC, et al. Auditory emotion recognition impairments in schizophrenia: relationship to acoustic features and cognition. Am J Psychiatry 2012 Apr;169(4):424-432 [FREE Full text] [doi: 10.1176/appi.ajp.2011.11081230] [Medline: 22362394]
49. Leitman DI, Laukka P, Juslin PN, Saccente E, Butler P, Javitt DC. Getting the cue: sensory contributions to auditory emotion recognition impairments in schizophrenia. Schizophr Bull 2010 May;36(3):545-556 [FREE Full text] [doi: 10.1093/schbul/sbn115] [Medline: 18791077]
50. Mendelson MJ, Aboud FE. Measuring friendship quality in late adolescents and young adults: McGill friendship questionnaires. Can J Behav Sci 1999;31(2):130-132. [doi: 10.1037/h0087080]
51. Sidner CL, Kidd CD, Lee C, Lesh N. Where to Look: A Study of Human-Robot Engagement. In: Proceedings of the 9th International Conference on Intelligent User Interfaces. 2004 Presented at: IUI'04; January 13-16, 2004; Funchal, Madeira, Portugal p. 78-84 URL: https://www.researchgate.net/publication/221608250_Where_to_look_A_study_of_human-robot_engagement [doi: 10.1145/964442.964458]
52. Hassenzahl M, Burmester M, Koller F. [Attrakdiff: ein fragebogen zur messung wahrgenommener hedonischer und pragmatischer qualität]. Mensch Comput 2003;57(1):187-196 [FREE Full text] [doi: 10.1007/978-3-322-80058-9_19]
53. Leite I, Pereira A, Mascarenhas S, Martinho C, Prada R, Paiva A. The influence of empathy in human–robot relations. Int J Hum-Comput St 2013 Mar;71(3):250-260. [doi: 10.1016/j.ijhcs.2012.09.005]
54. Mercer SW, Maxwell M, Heaney D, Watt GC. The consultation and relational empathy (CARE) measure: development and preliminary validation and reliability of an empathy-based consultation process measure. Fam Pract 2004 Dec;21(6):699-705. [doi: 10.1093/fampra/cmh621] [Medline: 15528286]
55. Bikker AP, Fitzpatrick B, Murphy D, Mercer SW. Measuring empathic, person-centred communication in primary care nurses: validity and reliability of the consultation and relational empathy (CARE) measure. BMC Fam Pract 2015 Oct 23;16:149 [FREE Full text] [doi: 10.1186/s12875-015-0374-y] [Medline: 26493072]
56. Wirtz M, Boecker M, Forkmann T, Neumann M. Evaluation of the 'consultation and relational empathy' (CARE) measure by means of rasch-analysis at the example of cancer patients. Patient Educ Couns 2011 Mar;82(3):298-306. [doi: 10.1016/j.pec.2010.12.009] [Medline: 21256692]

57. Law MM, Broadbent EA, Sollers JJ. A comparison of the cardiovascular effects of simulated and spontaneous laughter. Complement Ther Med 2018 Apr;37:103-109. [doi: 10.1016/j.ctim.2018.02.005] [Medline: 29609921]
58. Faul F, Erdfelder E, Lang A, Buchner A. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behav Res Methods 2007 May;39(2):175-191. [doi: 10.3758/BF03193146] [Medline: 17695343]
59. Bailenson JN, Aharoni E, Beall AC, Guadagno RE, Dimov A, Blascovich J. Comparing Behavioral and Self-Report Measures of Embodied Agents' Social Presence in Immersive Virtual Environments. In: Proceedings of the 7th Annual International Workshop on PRESENCE. 2004 Presented at: PRESENCE'04; October 13-15, 2004; Valencia, Spain URL: https://www.researchgate.net/publication/228996433_Comparing_behavioral_and_self-report_measures_of_embodied_agents%27_social_presence_in_immersive_virtual_environments
60. Bennett HJ. Humor in medicine. South Med J 2003 Dec;96(12):1257-1261. [doi: 10.1097/01.SMJ.0000066657.70073.14] [Medline: 14696878]
61. Sjoberg J, Araki K. Robots Make Things Funnier. In: Proceedings of the Annual Conference of the Japanese Society for Artificial Intelligence. 2008 Presented at: JSAI'08; June 11-13, 2008; Asahikawa, Japan p. 306-313 URL: https://link.springer.com/chapter/10.1007/978-3-642-00609-8_27#citeas [doi: 10.1007/978-3-642-00609-8_27]
62. Bertero D, Fung P. A Long Short-Term Memory Framework for Predicting Humor in Dialogues. In: Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. 2016 Presented at: NAACL-HLT'16; June 12-17, 2016; San Diego, California p. 12-17 URL: https://aclweb.org/anthology/papers/N/N16/N16-1016/ [doi: 10.18653/v1/N16-1016]
63. Robert Jr LP. Personality in the Human Robot Interaction Literature: A Review and Brief Critique. In: Proceedings of the 24th American Conference on Information Systems. 2018 Presented at: AMCIS'18; August 16-18, 2018; New Orleans, LA URL: https://www.researchgate.net/publication/325206897_Personality_in_the_Human_Robot_Interaction_Literature_A_Review_and_Brief_Critique

Abbreviations

ANOVA: analysis of variance
MR: mean rank

Edited by C Eichenberg; submitted 09.02.19; peer-reviewed by JA Sánchez Margallo, L Robert; comments to author 15.03.19; revised version received 06.05.19; accepted 17.07.19; published 04.10.19.

Please cite as:
Johanson DL, Ahn HS, MacDonald BA, Ahn BK, Lim J, Hwang E, Sutherland CJ, Broadbent E
The Effect of Robot Attentional Behaviors on User Perceptions and Behaviors in a Simulated Health Care Interaction: Randomized Controlled Trial
J Med Internet Res 2019;21(10):e13667
URL: https://www.jmir.org/2019/10/e13667
doi:10.2196/13667
PMID:31588904

Original Paper

A Protocol-Driven, Bedside Digital Conversational Agent to Support Nurse Teams and Mitigate Risks of Hospitalization in Older Adults: Case Control Pre-Post Study

Nicholas Bott[1,2], PhD; Sharon Wexler[3], RN, PhD; Lin Drury[3], RN, PhD; Chava Pollak[3], RN, MSN; Victor Wang[4], MS; Kathleen Scher[5], RN, EDD; Sharon Narducci[5], RN, DNP

[1]Clinical Excellence Research Center, Department of Medicine, Stanford University School of Medicine, Stanford, CA, United States

[2]Department of Psychology, PGSP-Stanford Consortium, Palo Alto, CA, United States

[3]Pace University, New York, NY, United States

[4]Care.Coach, Millbrae, CA, United States

[5]Jamaica Hospital Medical Center, New York, NY, United States

Corresponding Author:
Nicholas Bott, PhD
Clinical Excellence Research Center
Department of Medicine
Stanford University School of Medicine
75 Alta Road
Stanford, CA, 94305
United States
Phone: 1 650 814 9383
Email: nbott@stanford.edu

Abstract

Background: Hospitalized older adults often experience isolation and disorientation while receiving care, placing them at risk for many inpatient complications, including loneliness, depression, delirium, and falls. Embodied conversational agents (ECAs) are technological entities that can interact with people through spoken conversation. Some ECAs are also relational agents, which build and maintain socioemotional relationships with people across multiple interactions. This study utilized a novel form of relational ECA, provided by Care Coach (care.coach, inc): an animated animal avatar on a tablet device, monitored and controlled by live health advocates. The ECA implemented algorithm-based clinical protocols for hospitalized older adults, such as reorienting patients to mitigate delirium risk, eliciting toileting needs to prevent falls, and engaging patients in social interaction to facilitate social engagement. Previous pilot studies of the Care Coach avatar have demonstrated the ECA's usability and efficacy in home-dwelling older adults. Further study among hospitalized older adults in a larger experimental trial is needed to demonstrate its effectiveness.

Objective: The aim of the study was to examine the effect of a human-in-the-loop, protocol-driven relational ECA on loneliness, depression, delirium, and falls among diverse hospitalized older adults.

Methods: This was a clinical trial of 95 adults over the age of 65 years, hospitalized at an inner-city community hospital. Intervention participants received an avatar for the duration of their hospital stay; participants on a control unit received a daily 15-min visit from a nursing student. Measures of loneliness (3-item University of California, Los Angeles Loneliness Scale), depression (15-item Geriatric Depression Scale), and delirium (confusion assessment method) were administered upon study enrollment and before discharge.

Results: Participants who received the avatar during hospitalization had lower frequency of delirium at discharge (P<.001), reported fewer symptoms of loneliness (P=.01), and experienced fewer falls than control participants. There were no significant differences in self-reported depressive symptoms.

Conclusions: The study findings validate the use of human-in-the-loop, relational ECAs among diverse hospitalized older adults.

(J Med Internet Res 2019;21(10):e13440) doi:10.2196/13440

KEYWORDS

digital health; older adults; loneliness; delirium; falls; embodied conversational agent; chatbot; relational agent; information and communication technology

Introduction

Background

In 2014, 19% of Medicare beneficiaries had at least one inpatient stay covered under Medicare Part A [1]. In addition to the illness or injury requiring acute care, the unfamiliar and stressful environment of the hospital increases risk of loneliness, depression, delirium, and falls in these patients. Loneliness is a frequent occurrence in older adults, and it is a documented predictor of poor health. One in 3 older adults reports loneliness in the United States [2], and a recent American Association of Retired Persons (AARP) survey found that approximately 42.6 million older adults suffer from chronic loneliness [3]; research demonstrates that loneliness is linked to a variety of negative health outcomes, including high blood pressure [4], disability [5], functional decline [2], depression [6], and cognitive decline [7]. These comorbidities may consequently increase the need for health care, and these are linked to greater health care utilization [8].

For older adults, the hospital setting is also a precipitant for symptoms of depression and loss of control [9]. Depression, in turn, may stymie recovery and increase the length of hospital stays [10]. At the same time, a recent international study found that informal psychological support was inversely related with depressive symptoms in hospitalized older adults [10]. The challenge remains to find ways to incorporate low-cost interventions that provide psychological support within the fast-paced hospital setting.

In addition to loneliness and depression, delirium and falls represent adverse events for hospitalized older adults. Delirium is a serious but preventable condition associated with morbidity and mortality, occurring as frequently as 50% of all hospitalized older adults [11]. Owing to its limited recognition within inpatient settings, the presence of delirium is associated with inappropriate use of sedation, sitters, and restraints [12]. Furthermore, delirium contributes to falls, cognitive decline, disability, morbidity, and mortality, with estimated direct health care costs over US $164 billion a year in the United States [11,13,14].

Falls are a common adverse event associated with hospitalization, and fall risk increases with age. Delirium significantly increases the risk of falls in the hospital setting [11,13,14]. In the United States, all of these risks are greatest among patients with limited English comprehension, low health literacy, and socioeconomic disadvantage [15]. Evidence demonstrates that an individualized multifactorial approach to fall prevention, including alert wristbands, room signage, staff and patient education, footwear, toileting schedules, exercise, and movement alarms, is effective [16]. For patients with cognitive impairment, increased nursing surveillance is also an effective intervention [17].

In view of the relationship between these adverse events and morbidity, as well as decreased quality of life, increased health care usage, and subsequent cost, it is imperative to find interventions that can reduce the risk of these outcomes in hospitalized older adults. The Hospital Elder Life Program (HELP) is a multidomain nonpharmacological intervention that has been shown to be effective in mitigating the risk of these adverse events. A meta-analysis of 14 interventional studies demonstrated efficacy in reducing incidence of delirium [18], falls [18], and functional status [18]. Separately, HELP has also been shown to improve symptoms of patient loneliness [19]. HELP volunteers deliver protocols targeting multiple risk factors, including orientation, mobilization, vision, hearing, hydration, nutrition, and sleep. HELP is used in more than 200 hospitals worldwide, and it serves as the gold-standard nonpharmacological intervention for the risk mitigation and management of delirium, cognitive and functional decline, falls, and 1:1 observation among older adults [18-20]. Although the HELP program is cost effective, it requires substantial training and monitoring of a large team of volunteers, which can serve as a barrier to adoption, especially in safety-net facilities with limited resources, serving high-risk patient populations [21-23]. To date, research has not determined whether human-driven virtual companionship can approximate the outcomes of physical volunteers for hospitalized older adults.

The health technology literature is abundant and continues to grow; nevertheless, there is a dearth of information on the use of technology as a social relational agent for older adults. Embodied conversational agents (ECAs) are technological entities that can interact with people through spoken conversation. Some ECAs also function as relational agents and are designed to leverage these conversational interactions to build and maintain social-emotional relationships with people [24]. The Care Coach avatar is a relational ECA that interacts with patients primarily through speech, along with its own visual appearance, digital pictures, and audio or music that can play through the tablet speaker in the same way as a HELP volunteer would in vivo. The Care Coach avatar is powered by a human-in-the-loop software system, so-called because of a round-the-clock team of live health advocates monitoring and controlling each avatar. As a result, avatar conversations with patients and care teams are conducted through natural speech, allowing complex patients with cognitive and functional impairments to be engaged effectively, regardless of their technical abilities or inclination. Severely hearing-impaired patients may even understand the avatar by reading the captions displayed above the avatar's head on the tablet screen.

Existing studies have generally evaluated the ease of use and acceptability of health technology, and they have specifically evaluated social agents as a tool for decreasing loneliness and social isolation in older adults [25-29]. Khosravi et al conducted a systematic review to synthesize studies investigating the role of technology in addressing loneliness and social isolation in older adults [29]. These studies included different technological

modalities offering various means of engagement, including computer/internet use, robotics, video games, personal reminder information and social management systems, social networking sites, tele-care, and 3D virtual environments. Efficacy data included reduced loneliness, development and maintenance of social relationships, increased independent living, decreased social isolation, companionship, cognitive stimulation, and entertainment [29].

In a pilot study of home-dwelling older adults, Chi et al provided a qualitative perspective on utility and comfort with the Care Coach avatar presented in this paper [26]. Participants reported positive results in terms of companionship, social support, and health information. However, some concerns included internet connectivity, privacy, and cost. In addition, some participants reported variations in the quality of the avatar's conversations, associated with differing staff members operating the avatar, highlighting the importance of the human aspect of this technological experience. Demiris et al conducted a study utilizing the Care Coach platform in home-dwelling older adults with mild cognitive impairment and reported improvements in cognition and social support, along with decreases in depressive symptoms [30].

Objectives

The results of the previous pilot study demonstrated the feasibility of the Care Coach platform to improve clinically meaningful aspects of patient care in the home. Given the additional risks of delirium and falls, in addition to loneliness and depression, that hospitalized older adults experience during hospitalization, a less resource-intensive deployment of the multidomain interventions, included within the HELP program through a relational ECA, may provide a meaningful and scalable method for mitigating adverse events. This study investigated the efficacy of a human-in-the-loop, protocol-driven relational ECA, compared with control participants among hospitalized older adults on self-reported measures of loneliness, depression, and clinician-reported occurrence of delirium and falls. We hypothesized that the use of such an intervention would decrease self-reported loneliness and depression, as well as reduce the occurrence of delirium and falls during hospitalization.

Methods

Experimental Intervention

The relational ECA software platform utilized in the study was provided by Care Coach, a private company based in California, and it comprises an internet-based communication system, designed to provide 24-hour psychosocial and health care support for patients through an integrative, person-centered approach [31]. A relational ECA serves as the patient-facing interface, appearing as an animated dog or cat on the display screen of a tablet computing device (see Figure 1). The digital animal responds to touch and demonstrates emotions appropriate to the conversation or other interaction state, including facial expressions, bodily reaction to touch and petting, heart symbols, tears, sleeping, and snoring. The Care Coach platform represents a real-time fusion of human and software intelligence, powered by a team of live health advocates who see, hear, and speak with each patient through the avatar of the digital animal. Health advocates are available 24 hours a day, and they are guided by software algorithms to implement clinical protocols. The use of a nonhuman avatar plays several important roles. First, previous studies have found unique clinical benefits associated with nonhuman avatar relational agents [32]. Second, an avatar relational agent provides greater continuity of care across the duration of the hospital stay, avoiding coverage gaps and challenges with establishing and re-establishing a care relationship throughout the duration of the hospital stay. Finally, the use of a nonhuman avatar provides a clear distinction in role for patients. The avatar relational agent is not a medical professional that provides clinical information, and the use of a nonhuman avatar (as opposed to a nurse of clinician avatar) visually clarifies this distinction.

Using a proprietary Web-based work interface, health advocates can sequentially monitor and engage 12 or more patients, and they can simultaneously monitor and engage up to 2 patients. Health advocates observe and listen through the audiovisual feed from the avatar device, communicating in real time with each patient by sending text commands, which are converted into the avatar's voice using a text-to-speech engine. Thus, the health advocates contribute their human abilities for natural language understanding, in both English and Spanish, and sociable, compassionate, and conversational responses to help patients build personal relationships with the avatar. Their human abilities are augmented through a software-driven system embedded into the work interface, which uses branching logic and prescripted conversational content to guide health advocates through evidence-based intervention protocols, such as cognitive engagement or assessment, reorientation, toileting checks, and fall prevention (see Figure 2), alerting the nurse station on the hospital unit by phone of any issues potentially requiring immediate action.

In this study, the avatars were programmed with specific protocols, repeating for each patient every 2 days (see Multimedia Appendix 1). For example, the Care Coach system was programmed to proactively ask about comfort and need of bathroom use, seeking assistance on behalf of the patient by calling the nursing station. Studies show that approximately half of all falls in the hospital occur either in the bathroom or on the way to the bathroom [33,34]. In addition to scheduled protocols, each avatar provided engagement on an unscheduled, informal basis through relationship-building conversations about participant interests, family, and news, as well as through showing pictures, playing music, guided meditation/relaxation tracks, and engaging the patient in brain games.

Figure 1. The Care Coach avatar system design (top). An avatar encourages a patient to take his medications (bottom).

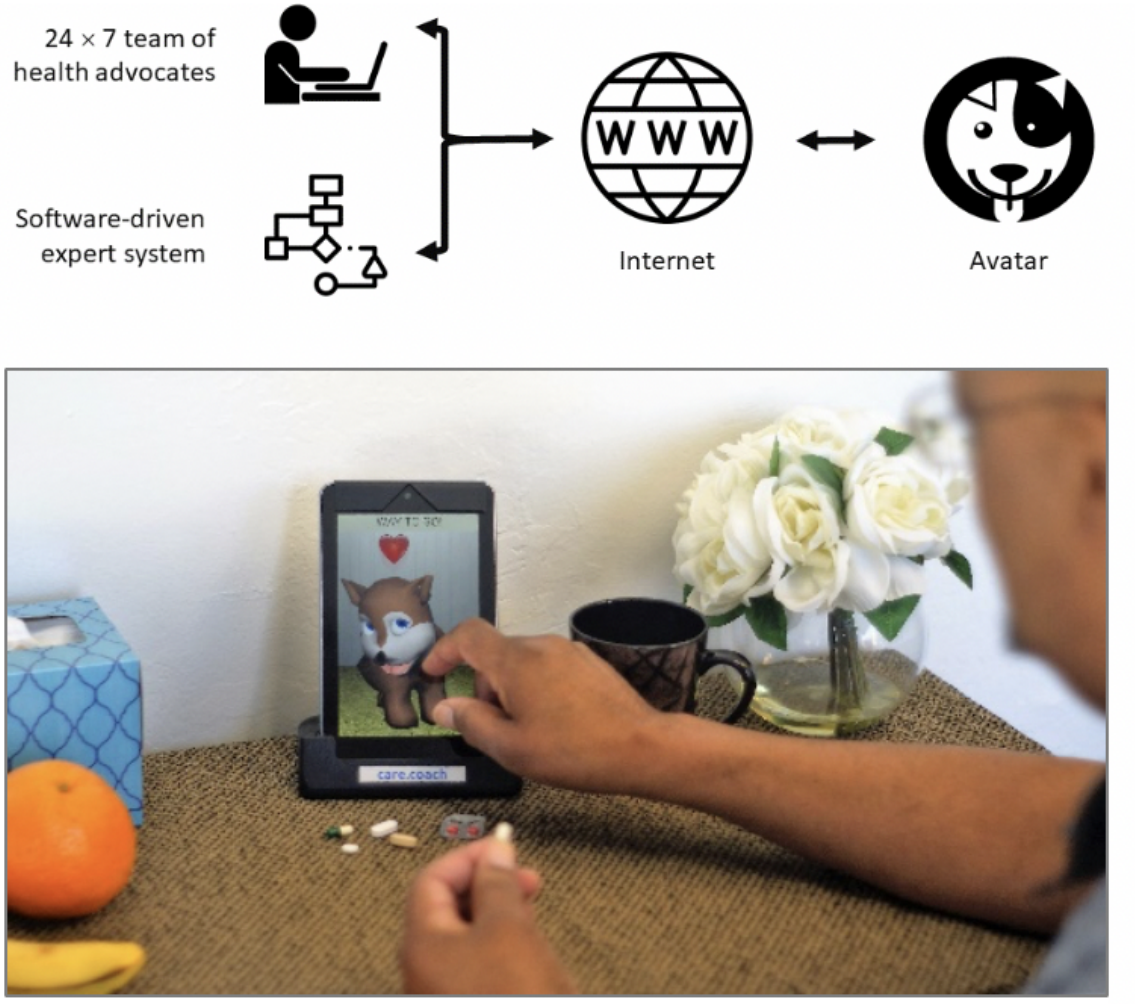
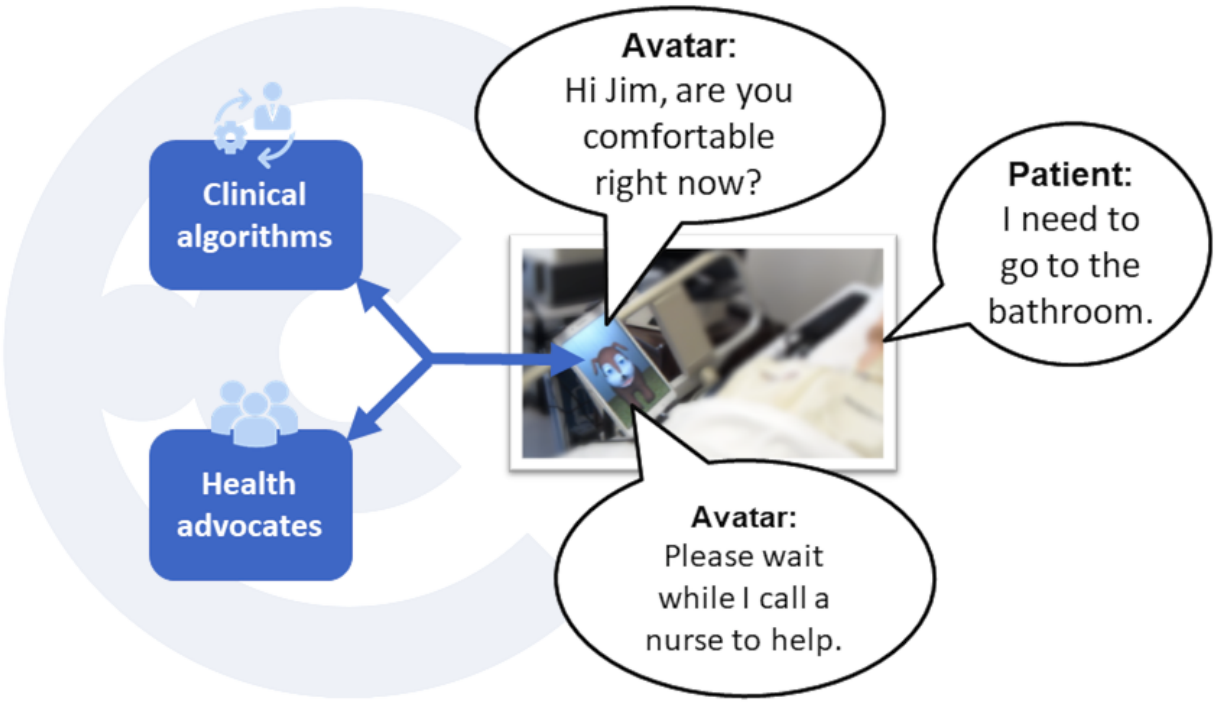


Figure 2. An example of a simple fall prevention protocol.



As a relational ECA, the health advocate is able to wake up the avatar with an accompanying *wake noise* whenever the audiovisual stream is started and visually put the avatar to sleep whenever the audiovisual stream stops. As the Care Coach platform utilizes a human-in-the-loop health advocate, all patient information is treated as protected health information, and all health advocates receive Health Insurance Portability and Accountability Act training. Moreover, patients may ask the avatar for privacy at any time, as long as such a request is consistent with the safety monitoring requirements of the health care provider or proxy.

Study Design

This clinical study used a case control quasi-experimental pre-post design. Participants were recruited from 3 medical-surgical units in a 600+ bed community hospital. Units were similar in size (34-37 beds), average daily census (32 patients), and average length of stay (3-6 days). One unit served as the control unit, with 2 units serving as intervention units. Patients were enrolled and followed by a team of research assistants (RAs), comprising undergraduate and graduate nursing and computer science students. The RAs were trained, and they followed a structured protocol and scripts. Project management, as well as scheduling of RAs, was conducted by a graduate nursing student.

Setting

The study was conducted in an urban community hospital in New York City, which targets the underserved. The hospital is in the borough of Queens, one of the most ethnically diverse urban areas in the world [35]. The majority of the borough's population identifies as nonwhite, with strong representation from black, Asian, and Hispanic races [36]. Per capita income is US $28,814, with 19% of adults reporting less than high school education, and 11% of adults reporting lack of health insurance [36]. The hospital has 408 beds and approximately 120,000 emergency department visits annually [37].

Participants and Procedures

Initial screening of potential participants was conducted by the nurse managers of the respective nursing units during daily rounds. Inclusion criteria were patients over the age of 65, admitted to 1 of the study units, who could give informed consent or had a proxy who could provide consent. Exclusion criteria were patients who could not provide informed consent with no proxy, as well as patients who wandered, demonstrated aggressive combative behavior with intent to harm self or others, or were experiencing alcohol or drug withdrawal. Patients who were noncommunicative or did not speak English were also excluded from the study. Eligible patients were referred to an RA. RAs discussed the study, obtained informed consent, and administered the enrollment instruments, which comprised a demographic questionnaire and assessment instruments (detailed below). The enrollment visit lasted approximately 15 min per participant.

A total of 2 units were utilized for the experimental intervention, with a third unit utilized as a control unit. All patients recruited from either of the intervention units were enrolled into the intervention group. All patients recruited from the control unit were enrolled into the control group. Owing to the pace of recruitment, enrollment in the second experimental unit began approximately 45 days after the first experimental unit. Enrolled patients on the intervention units selected either a dog or a cat avatar. RAs demonstrated how the avatar works and verified that each subject was comfortable with the device before leaving the bedside. RAs visited each subject daily throughout the hospitalization to confirm that each avatar was functioning and within view of the patient. Upon 24-hour discharge notification, RAs administered the postassessment instruments. Subjects discharged home or to another facility had the option of taking their avatar with them. Home-based data were collected on those participants utilizing the avatar postdischarge, which will be the focus of a separate follow-up study. Relational agent theory [38] suggests that conversational variety is essential to sustaining interaction between the subject and the avatar. Extending the period over which a protocol set is repeated promotes sustained engagement; therefore, rather than daily repetition, a 2-day rotation period was designed on the basis of an expected patient length of stay of approximately 4 days. Enrolled patients on the control unit received the same preassessments as the intervention participants. RAs visited each control participant daily throughout the hospitalization to provide a "dose" of RA contact, similar to that received by the intervention participants. RAs used a rounding script (see Multimedia Appendix 2) for both control and intervention patients, to ensure consistency among visits and equalize the influence of human contact on the outcomes being measured. Upon 24-hour discharge notification, RAs administered postassessment instruments to the control patients. Throughout the study, RAs recorded observations on the subjects' responses to the avatars, using the observation checklist. All procedures were approved by the Jamaica Hospital and PACE University Institutional Review Boards.

Measures

Demographic variables, including age, gender, education, race/ethnicity, and living arrangements, were collected upon enrollment. Discharge destination and living arrangements were collected up to 24 hours before discharge. Delirium, loneliness, and depression were assessed using the instruments detailed below. All instruments were administered at the time of enrollment into the study and upon discharge.

Delirium

The confusion assessment method (CAM) was utilized to screen for delirium [39]. The CAM includes an instrument and diagnostic algorithm for identification of delirium, assessing the presence, severity, and fluctuation of 9 delirium features, including the following: acute onset, inattention, disorganized thinking, altered level of consciousness, disorientation, memory impairment, perceptual disturbances, psychomotor agitation or retardation, and altered sleep-wake cycle. The CAM diagnostic algorithm is based on 4 cardinal features of delirium: (1) acute onset and fluctuating course, (2) inattention, (3) disorganized thinking, and (4) altered level of consciousness. A diagnosis of delirium utilizing the CAM requires the presence of features 1, 2, and either 3 or 4. The CAM has been used extensively for assessment of delirium among hospitalized older adults, and it

has demonstrated excellent sensitivity and specificity among large and small samples of older adults [39].

Loneliness

Loneliness was measured using the University of California Los Angeles (UCLA) Loneliness Scale Short Form [40]. Participants rated themselves on a scale of "1 to 3" on the 3 items ("hardly ever"=1, "some of the time"=2, and "often"=3), with a total score ranging from 3 to 9. Higher scores indicate more loneliness. The short form has demonstrated reliability with the long form of the measure [40].

Depression

Depression was measured using the Geriatric Depression Scale (GDS) 15-item short form [41]. Of the 15 items, 10 indicate the presence of depression when answered positively, whereas the rest indicate depression when answered negatively. Scores from 0 to 4 are considered normal, depending on age, education, and complaints; scores from 5 to 8 indicate mild depression, 9 to 11 indicate moderate depression, and 12 to 15 indicate severe depression. The GDS has been used extensively to measure depressive symptoms among older adults [42].

Falls

Data on falls were obtained from existing unit-based quality improvement data from the hospital. Incidence data are recorded quarterly in hospital reports and presented as a ratio in the form of falls per 1000 patient days [43]. The fall incidence data for this study were collected over 2 full calendar quarters, comprising the 3-month period, preceding any intervention, and the following 3-month period, during which the intervention was applied. The control unit did not receive any avatars in either 3-month period. One intervention unit received avatars and had all its enrolled patients treated as subjects during the 3-month intervention period. Owing to recruitment cadence, the additional intervention unit started the 3-month intervention period approximately 45 days into the calendar quarter. As a result, the falls data on the second intervention unit represent a mixed unit, with the first half of the quarterly falls data not including participants receiving the intervention. Only summary data on quarterly fall rates by unit and time period were available. As a result, these data were interpreted in relationship to baseline data and in relationship to the national average [44] (see Multimedia Appendix 3).

Power Analysis

Estimated sample size for this study utilized published results of the CAM, UCLA 3-item loneliness scale, and GDS. Medium-to-large effect sizes have been observed in studies utilizing the CAM, UCLA 3-item loneliness scale, and the GDS among hospitalized older adults [45,46]. Sample size calculation utilized an expected medium effect size (Cohen d=0.60), with alpha set at .05, and assumed power of 80% for analysis of variance yielded an expected sample size of 90. The study sample is powered at greater than 80% to detect a medium effect between groups.

Data Analysis

Two-tailed Student t test and chi-square cross tabulations were used for descriptive analysis of participants. McNemar test was utilized to compare change in frequency of delirium within intervention and control groups. Analysis of covariance was used to analyze the differences among intervention and control group means on measures of loneliness and depression. Sex, baseline self-reported loneliness, and baseline MiniCog score were included as covariates in the loneliness analysis. Sex, baseline self-reported depressive symptoms, and baseline MiniCog score were included in the depression analysis. Quarterly fall rates for intervention and control units were examined descriptively in relationship to baseline data and to the national average. The analyses for this study were done using the SPSS version 25 (SPSS Inc).

Results

Descriptive Analysis

A total of 95 participants were included in the analyses (intervention group n=41; control group n=54). There were no differences between the control and intervention groups on age, race, place of residence, discharge location, or language (P>.05). The intervention group (68%, 28/41 female) included more female participants than the control group (44%, 24/54 female; P=.02). Across groups, participants were predominantly female, English-speaking African Americans with a mean age of 76 years. They resided at home and were discharged to home post hospital stay (Table 1).

On average, the avatars checked in with participants in the intervention group (a health advocate started the audio/video stream, visually waking up the avatar) 71.3 times per day per patient. Avatars engaged intervention participants for 61 min per day (including average use of 11.5 images or audio files) and completed, on average, 6.5 protocol-driven tasks per day (Table 2).

Delirium

The presence of delirium at enrollment and discharge was analyzed within each group. McNemar tests found a significant reduction in delirium presence from enrollment to discharge in the intervention group (P<.001). There was no change in frequency of delirium presence within the control group (P=.25; Table 3).

Loneliness

A general linear model examining admission and discharge endorsement on the UCLA Loneliness Scale indicated that participants with avatars experienced a decrease in loneliness (P=.01) compared with participants in the control group (see Table 3).

Depression

A general linear model examining admission and discharge endorsement of depression symptomology on the GDS between groups showed no statistically significant difference in depression between participants with and without avatars (see Table 3).

Falls

Quarterly unit-based quality improvement data on incidence of falls per 1000 patient days were examined across each of the 3

units. Falls rate was reduced by 33% on the avatar unit, with partial data collection. Falls rate was reduced by 82% on the avatar unit, with complete data collection. Frequency of falls increased on the control unit by 86%.

Table 1. Patient characteristics by study group.

Characteristics	Intervention (n=41)	Control (n=54)	*P* value
Age (years), mean (SD)	76.88 (8.85)	76.22 (8.05)	.70
Race, n (%)			.68
White	10 (24)	14 (26)	
African American	22 (54)	22 (41)	
Asian/Pacific Islander	5 (12)	12 (22)	
Hispanic	3 (7)	5 (9)	
Other	1 (2)	1 (2)	
Sex, n (%)			.02[a]
Male	13 (32)	30 (56)	
Female	28 (68)	24 (44)	
Residence before admission, n (%)			.16
Home	38 (93)	46 (85)	
Nursing home	2 (5)	8 (15)	
Other	1 (2)	0 (0)	
Discharge location, n (%)			.59
Homeless	1 (2)	0 (0)	
Home	32 (78)	45 (83)	
Nursing home	5 (12)	7 (13)	
Short-term rehabilitation	1 (3)	0 (0)	
Other	2 (5)	2 (4)	
English as a second language, n (%)			.34
No	25 (61)	28 (52)	
Yes	16 (39)	26 (48)	

[a]*P*<.05.

Table 2. Patient engagement data (n=41).

Engagement metric	Mean (SD) per day
Number of check-ins	71.30 (7.46)
Observational and engagement time (min)	61.00 (40.61)
Media files used	11.50 (9.04)
Protocol tasks completed	6.5 (6.03)

Table 3. Outcome results.

Outcome variable	Intervention pre (n=41)	Intervention post (n=41)	Control pre (n=54)	Control post (n=54)	*P* value	Partial eta squared
Delirium, n/N (%)	12/29 (41)	1/40 (3)	6/48 (13)	3/51 (6)	<.001/.25	__[a]
Loneliness, mean (SD)	4.98 (2.17)	3.76 (1.53)	4.72 (1.74)	4.35 (1.70)	.01	0.07
Depression, mean (SD)	4.2 (3.2)	4.02 (2.99)	4.19 (3.5)	3.87 (2.99)	.46	0.006

[a]Not applicable.

Discussion

Principal Findings

This study investigated the efficacy of a human-in-the-loop, relational ECA in the form of a tablet-based virtual service animal avatar. Measures of delirium, loneliness, depression, and falls were investigated in a case-control study of 95 hospitalized adults over the age of 65 at an inner-city community hospital. Analysis of admission and discharge data indicated that intervention participants experienced lower frequency of delirium at discharge (P<.001) and a reduction in symptoms of loneliness (P=.01). Quarterly unit falls per 1000 patient days indicated that falls on the control unit increased by 86%. Falls on the intervention unit with delayed data collection were reduced by 33%. Falls on the intervention unit with complete data collection were reduced by 82%. There were no differences between groups in self-reported depressive symptoms. These results are consistent with previous research investigating devices that are used for a variety of needs specific to older adults, including safety and fall prevention, assistance with physical tasks that support activities of daily living, prevention of hospitalization, and social connectedness [28,29].

The avatar intervention resulted in a 91% reduction in delirium at discharge, with 11 participants meeting criteria for delirium with the CAM upon admission and 1 meeting criteria for delirium upon discharge. Although no individuals in either the control or intervention units acquired delirium after beginning the study, the Care Coach intervention's ability to help resolve delirium within a short period of days is comparable with data reported on HELP intervention among hospitalized older adults [21].

With respect to falls, the reduction observed across intervention units is likely attributable to the nature of preprogrammed avatar protocols assisting with nursing calls to assist with mobility and toileting. The reduction in falls seen across the study duration provides further validation for the use of a human-in-the-loop, tablet-based intervention to provide a means for hospitalized older adults to communicate with health care providers. This is important, given the role that alarm fatigue can play in health care service delivery [47]. The reason for the increase in falls rate on the control unit is not known.

There are currently 12 million Americans over the age 65 years living alone [48]. In a survey published by investigating the frequency of loneliness among 3000 older adults, 35% were categorized as being lonely [49]. The effect of the avatar intervention on symptoms of loneliness suggests that this form of human-in-the-loop virtual engagement is an effective means for combating this issue while adults are hospitalized. This is noteworthy, given that the participants without avatars received personal visits from RAs for approximately 15 min each hospital day. Contrary to our hypothesis, the avatar intervention did not demonstrate reduction in depressive symptoms when compared with the control group. This may be attributable to the tenacity of depression among older adults and the relatively short exposure to the avatars associated with the hospital stay (averaging 4 days). It may also be because of the relatively minimal number of depressive symptoms endorsed by participants across both groups at baseline.

An important implication of this study is the feasibility of larger deployment. As a low-cost and sustainable intervention, the marginal resources required to deploy such a human-in-the-loop avatar intervention model would comprise marginal staffing effort and minimal technology components, including health advocate staffing (which can take advantage of an abundant remote workforce), certified nursing assistant's effort to setup/sanitize the device (<US $1/patient day), wear and tear on the tablet device and related hardware (< US $1/patient day), and internet connection and server costs (approximately US $1/patient day on cellular data or approximately US $0/patient day with a reliable Wi-Fi network). Furthermore, participant learning curve is minimal, as the Care Coach technology platform does not require previous information technology fluency or use. Participants with mobility issues can engage with this technology, as the patient only needs to talk with the avatar, with no manual effort required. The digital animal avatar also solves such issues as the cost of caring for pets and potential animal-related allergies. Finally, preliminary evidence presented in this paper suggests the avatar may represent a superior intervention for risk reduction of delirium and falls over that of a patient sitter. Current evidence does not support the efficacy of patient sitters for reducing falls, and although there is a lack of evidence regarding patient sitters and delirium, a passive clinical care member in a patient's room would likely have minimal impact on delirium risk.

Additional research is needed to establish the long-term effects of an avatar on the well-being of patients and their ability to provide self-care. Patients who utilized the avatar during this study experienced less delirium and fell less than those patients who did not utilize the avatar. Given that both delirium and falls have significant morbidity and mortality in older adults and represent a major financial burden to the health care system, the use of this technology could potentially improve outcomes for hospitalized older adults. The cost of this intervention is significantly less than the cost of a fall (the average hospital cost for a fall injury is over US $30,000) [50] or the potential cost of increased length of stay because of delirium, which is an average increase of 7.78 days per case of incident delirium [51].

Limitations

This clinical study was limited by several factors. Although the sample size was sufficient, the intervention arm included convenience sampling and a potential bias in the referral process by the nurse managers. Given the deployment of this intervention in a real-world point-of-care environment, potential differences in unit care model and participant exposure to care across units could not be controlled. In addition, only patients who were able to understand and speak English were enrolled. In a very ethnically diverse setting, with many non-English speaking patients, many potential participants were excluded. However, the study location and population provided for the recruitment of a diverse population. Falls data were only available quarterly by unit. This limited investigation of falls data to percentage changes within units. The slower recruitment

cadence on the second experimental unit in combination with quarterly falls data resulted in partial data on falls for this unit. Finally, more granular data on individual patients were not available, which prevented investigation into potential effects related to specific patient medical comorbidities.

Conclusions

In conclusion, this study provides additional quantitative support to the clinical efficacy of the Care Coach avatar platform for hospitalized older adults. As a human-in-the-loop, tablet-based intervention, this technology provides a novel, scalable solution to mitigate the risks of delirium, loneliness, and falls among diverse hospitalized older adults. Future study of the Care Coach platform, including a randomized control trial (NCT03832192), will allow the replicability of these findings to be tested in a well-characterized and randomized patient population.

Acknowledgments

This research was supported by a grant from New York City Economic Development Corporation and Health 2.0.

Conflicts of Interest

VW is the CEO of and owns equity in care.coach. None of the other authors own equity in care.coach. SW and LD receive support from an ongoing National Institute of Nursing Research grant in conjunction with care.coach.

Multimedia Appendix 1
Avatar protocols.
[PDF File (Adobe PDF File), 78 KB - jmir_v21i10e13440_app1.pdf]

Multimedia Appendix 2
Rounding script.
[PDF File (Adobe PDF File), 60 KB - jmir_v21i10e13440_app2.pdf]

Multimedia Appendix 3
Quarterly fall rates by hospital unit and time periods.
[PNG File , 483 KB - jmir_v21i10e13440_app3.png]

References

1. MedPAC. 2016. Data Book: Health Care Spending and the Medicare Program URL: http://www.medpac.gov/docs/default-source/data-book/jun19_databook_entirereport_sec.pdf?sfvrsn=0 [accessed 2019-09-04]
2. Perissinotto CM, Cenzer IS, Covinsky KE. Loneliness in older persons: a predictor of functional decline and death. Arch Intern Med 2012 Jul 23;172(14):1078-1083 [FREE Full text] [doi: 10.1001/archinternmed.2012.1993] [Medline: 22710744]
3. National Resource Center on Nutrition and Aging. 2017. Loneliness among Older Adults: A National Survey of Adults 45+ URL: https://nutritionandaging.org/loneliness-among-older-adults-a-national-survey-of-adults-45/ [accessed 2017-09-04]
4. Hawkley LC, Thisted RA, Masi CM, Cacioppo JT. Loneliness predicts increased blood pressure: 5-year cross-lagged analyses in middle-aged and older adults. Psychol Aging 2010 Mar;25(1):132-141 [FREE Full text] [doi: 10.1037/a0017805] [Medline: 20230134]
5. Lund R, Nilsson CJ, Avlund K. Can the higher risk of disability onset among older people who live alone be alleviated by strong social relations? A longitudinal study of non-disabled men and women. Age Ageing 2010 May;39(3):319-326. [doi: 10.1093/ageing/afq020] [Medline: 20208073]
6. Cacioppo JT, Hughes ME, Waite LJ, Hawkley LC, Thisted RA. Loneliness as a specific risk factor for depressive symptoms: cross-sectional and longitudinal analyses. Psychol Aging 2006 Mar;21(1):140-151. [doi: 10.1037/0882-7974.21.1.140] [Medline: 16594799]
7. James BD, Wilson RS, Barnes LL, Bennett DA. Late-life social activity and cognitive decline in old age. J Int Neuropsychol Soc 2011 Nov;17(6):998-1005 [FREE Full text] [doi: 10.1017/S1355617711000531] [Medline: 22040898]
8. Taube E, Kristensson J, Sandberg M, Midlöv P, Jakobsson U. Loneliness and health care consumption among older people. Scand J Caring Sci 2015 Sep;29(3):435-443. [doi: 10.1111/scs.12147] [Medline: 24826811]
9. Bridges J, Flatley M, Meyer J. Older people's and relatives' experiences in acute care settings: systematic review and synthesis of qualitative studies. Int J Nurs Stud 2010 Jan;47(1):89-107. [doi: 10.1016/j.ijnurstu.2009.09.009] [Medline: 19854441]
10. Gur-Yaish N, Zisberg A, Sinoff G, Shadmi E. Effects of instrumental and psychological support on levels of depressive symptoms for hospitalized older adults. Aging Ment Health 2013;17(5):646-653. [doi: 10.1080/13607863.2012.758234] [Medline: 23330681]

11. Labella AM, Merel SE, Phelan EA. Ten ways to improve the care of elderly patients in the hospital. J Hosp Med 2011;6(6):351-357. [doi: 10.1002/jhm.900] [Medline: 21834118]
12. Adams J, Kaplow R. A sitter-reduction program in an acute health care system. Nurs Econ 2013;31(2):83-89. [Medline: 23691749]
13. Inouye SK, Westendorp RG, Saczynski JS. Delirium in elderly people. Lancet 2014 Mar 8;383(9920):911-922 [FREE Full text] [doi: 10.1016/S0140-6736(13)60688-1] [Medline: 23992774]
14. Leslie DL, Inouye SK. The importance of delirium: economic and societal costs. J Am Geriatr Soc 2011 Nov;59(Suppl 2):S241-S243 [FREE Full text] [doi: 10.1111/j.1532-5415.2011.03671.x] [Medline: 22091567]
15. Tow A, Holtzer R, Wang C, Sharan A, Kim SJ, Gladstein A, et al. Cognitive reserve and postoperative delirium in older adults. J Am Geriatr Soc 2016 Jun;64(6):1341-1346 [FREE Full text] [doi: 10.1111/jgs.14130] [Medline: 27321616]
16. Oliver D, Healey F, Haines TP. Preventing falls and fall-related injuries in hospitals. Clin Geriatr Med 2010 Nov;26(4):645-692. [doi: 10.1016/j.cger.2010.06.005] [Medline: 20934615]
17. Mazur K, Wilczyński K, Szewieczek J. Geriatric falls in the context of a hospital fall prevention program: delirium, low body mass index, and other risk factors. Clin Interv Aging 2016;11:1253-1261 [FREE Full text] [doi: 10.2147/CIA.S115755] [Medline: 27695303]
18. Hshieh TT, Yue J, Oh E, Puelle M, Dowal S, Travison T, et al. Effectiveness of multicomponent nonpharmacological delirium interventions: a meta-analysis. JAMA Intern Med 2015 Apr;175(4):512-520 [FREE Full text] [doi: 10.1001/jamainternmed.2014.7779] [Medline: 25643002]
19. Steunenberg B, van der Mast R, Strijbos MJ, Inouye SK, Schuurmans MJ. How trained volunteers can improve the quality of hospital care for older patients. A qualitative evaluation within the hospital elder life program (HELP). Geriatr Nurs 2016;37(6):458-463. [doi: 10.1016/j.gerinurse.2016.06.014] [Medline: 27471215]
20. Inouye SK, Bogardus ST, Baker DI, Leo-Summers L, Cooney LM. The hospital elder life program: a model of care to prevent cognitive and functional decline in older hospitalized patients. Hospital elder life program. J Am Geriatr Soc 2000 Dec;48(12):1697-1706. [doi: 10.1111/j.1532-5415.2000.tb03885.x] [Medline: 11129764]
21. Rubin FH, Williams JT, Lescisin DA, Mook WJ, Hassan S, Inouye SK. Replicating the hospital elder life program in a community hospital and demonstrating effectiveness using quality improvement methodology. J Am Geriatr Soc 2006 Jun;54(6):969-974. [doi: 10.1111/j.1532-5415.2006.00744.x] [Medline: 16776794]
22. Rubin FH, Neal K, Fenlon K, Hassan S, Inouye SK. Sustainability and scalability of the hospital elder life program at a community hospital. J Am Geriatr Soc 2011 Feb;59(2):359-365 [FREE Full text] [doi: 10.1111/j.1532-5415.2010.03243.x] [Medline: 21314654]
23. Caplan GA, Harper EL. Recruitment of volunteers to improve vitality in the elderly: the REVIVE study. Intern Med J 2007 Feb;37(2):95-100. [doi: 10.1111/j.1445-5994.2007.01265.x] [Medline: 17229251]
24. Bickmore TW, Caruso L, Clough-Gorr K, Heeren T. 'It's just like you talk to a friend' relational agents for older adults. Interact Comput 2005 Dec;17(6):711-735. [doi: 10.1016/j.intcom.2005.09.002]
25. Beuscher LM, Fan J, Sarkar N, Dietrich MS, Newhouse PA, Miller KF, et al. Socially assistive robots: measuring older adults' perceptions. J Gerontol Nurs 2017 Dec 1;43(12):35-43 [FREE Full text] [doi: 10.3928/00989134-20170707-04] [Medline: 28700074]
26. Chi NC, Sparks O, Lin SY, Lazar A, Thompson HJ, Demiris G. Pilot testing a digital pet avatar for older adults. Geriatr Nurs 2017;38(6):542-547 [FREE Full text] [doi: 10.1016/j.gerinurse.2017.04.002] [Medline: 28479082]
27. Liang A, Piroth I, Robinson H, MacDonald B, Fisher M, Nater UM, et al. A pilot randomized trial of a companion robot for people with dementia living in the community. J Am Med Dir Assoc 2017 Oct 1;18(10):871-878. [doi: 10.1016/j.jamda.2017.05.019] [Medline: 28668664]
28. García-Soler Á, Facal D, Díaz-Orueta U, Pigini L, Blasi L, Qiu R. Inclusion of service robots in the daily lives of frail older users: a step-by-step definition procedure on users' requirements. Arch Gerontol Geriatr 2018 Jan;74:191-196. [doi: 10.1016/j.archger.2017.10.024] [Medline: 29128788]
29. Khosravi P, Rezvani A, Wiewiora A. The impact of technology on older adults' social isolation. Comput Hum Behav 2016 Oct;63:594-603. [doi: 10.1016/j.chb.2016.05.092]
30. Demiris G, Thompson HJ, Lazar A, Lin SY. Evaluation of a digital companion for older adults with mild cognitive impairment. AMIA Annu Symp Proc 2016;2016:496-503 [FREE Full text] [Medline: 28269845]
31. Wang V, Osborne T. Social robots and other relational agents to improve patient care. In: Chau D, editor. Using Technology to Improve Care of Older Adults. New York: Springer Publishing Company; 2018.
32. Burton A. Dolphins, dogs, and robot seals for the treatment of neurological disease. Lancet Neurol 2013 Sep;12(9):851-852. [doi: 10.1016/S1474-4422(13)70206-0] [Medline: 23948175]
33. Morgan VR, Mathison JH, Rice JC, Clemmer DI. Hospital falls: a persistent problem. Am J Public Health 1985 Jul;75(7):775-777. [doi: 10.2105/ajph.75.7.775] [Medline: 4003655]
34. Tzeng HM. Understanding the prevalence of inpatient falls associated with toileting in adult acute care settings. J Nurs Care Qual 2010;25(1):22-30. [doi: 10.1097/NCQ.0b013e3181afa321] [Medline: 19553863]
35. The Official Website of New York State. Queens: Overview URL: https://www.ny.gov/counties/queens [accessed 2019-01-18]

36. United States Census Bureau. Quick Facts: Queens County (Queens Borough), New York URL: https://www.census.gov/quickfacts/fact/table/queenscountyqueensboroughnewyork/POP645217 [accessed 2019-01-18]
37. Jamaica Hospital Medical Center. Hospital Facts URL: https://jamaicahospital.org/general-info/hospital-facts/ [accessed 2019-01-18]
38. Bickmore T, Gruber A, Picard R. Establishing the computer-patient working alliance in automated health behavior change interventions. Patient Educ Couns 2005 Oct;59(1):21-30. [doi: 10.1016/j.pec.2004.09.008] [Medline: 16198215]
39. Wei LA, Fearing MA, Sternberg EJ, Inouye SK. The confusion assessment method: a systematic review of current usage. J Am Geriatr Soc 2008 May;56(5):823-830 [FREE Full text] [doi: 10.1111/j.1532-5415.2008.01674.x] [Medline: 18384586]
40. Hughes ME, Waite LJ, Hawkley LC, Cacioppo JT. A short scale for measuring loneliness in large surveys: results from two population-based studies. Res Aging 2004;26(6):655-672 [FREE Full text] [doi: 10.1177/0164027504268574] [Medline: 18504506]
41. Yesavage JA, Sheikh JI. 9/geriatric depression scale (GDS). Clin Gerontol 2008 Oct 25;5(1-2):165-173. [doi: 10.1300/j018v05n01_09]
42. Marc LG, Raue PJ, Bruce ML. Screening performance of the 15-item geriatric depression scale in a diverse elderly home care population. Am J Geriatr Psychiatry 2008 Nov;16(11):914-921 [FREE Full text] [doi: 10.1097/JGP.0b013e318186bd67] [Medline: 18978252]
43. Agency for Healthcare Research and Quality. 2013. Preventing Falls in Hospitals: How do you measure fall rates and fall prevention practices? URL: https://www.ahrq.gov/professionals/systems/hospital/fallpxtoolkit/fallpxtk5.html [accessed 2019-09-04]
44. Chang SY, Chen WS, Teng T, Yeh CY, Yen HC. Fall risk program for oncology inpatients: addition of the 'traffic light' fall risk assessment tool. J Nurs Care Qual 2019;34(2):139-144. [doi: 10.1097/NCQ.0000000000000353] [Medline: 30198946]
45. Singh I, Subhan Z, Krishnan M, Edwards C, Okeke J. Loneliness among older people in hospitals: a comparative study between single rooms and multi-bedded wards to evaluate current health service within the same organisation. Gerontol Geriatr Res 2016;2(3):1015 [FREE Full text]
46. Koenig HG, Meador KG, Cohen HJ, Blazer DG. Depression in elderly hospitalized patients with medical illness. Arch Intern Med 1988 Sep;148(9):1929-1936. [doi: 10.1001/archinte.1988.00380090031009] [Medline: 3415405]
47. Vandenberg AE, van Beijnum B, Overdevest VG, Capezuti E, Johnson TM. US and Dutch nurse experiences with fall prevention technology within nursing home environment and workflow: a qualitative study. Geriatr Nurs 2017;38(4):276-282 [FREE Full text] [doi: 10.1016/j.gerinurse.2016.11.005] [Medline: 27956058]
48. Stepler R. Pew Social Trends. 2016. Smaller Share of Women Ages 65 and Older Are Living Alone URL: http://www.pewsocialtrends.org/2016/02/18/smaller-share-of-women-ages-65-and-older-are-living-alone/ [accessed 2019-09-04]
49. Anderson GO. AARP. 2010. Loneliness Among Older Adults: A National Survey of Adults 45+ URL: http://www.aarp.org/research/topics/life/info-2014/loneliness_2010.html [accessed 2018-12-24]
50. Burns ER, Stevens JA, Lee R. The direct costs of fatal and non-fatal falls among older adults - United States. J Safety Res 2016 Sep;58:99-103. [doi: 10.1016/j.jsr.2016.05.001] [Medline: 27620939]
51. McCusker J, Cole MG, Dendukuri N, Belzile E. Does delirium increase hospital stay? J Am Geriatr Soc 2003 Nov;51(11):1539-1546. [doi: 10.1046/j.1532-5415.2003.51509.x] [Medline: 14687382]

Abbreviations

CAM: confusion assessment method
ECA: embodied conversational agent
GDS: Geriatric Depression Scale
HELP: The Hospital Elder Life Program
RA: research assistant
UCLA: University of California Los Angeles

Edited by G Eysenbach; submitted 18.01.19; peer-reviewed by M Immonen, M Stein, S Ge; comments to author 25.04.19; revised version received 21.07.19; accepted 19.08.19; published 17.10.19.

Please cite as:
Bott N, Wexler S, Drury L, Pollak C, Wang V, Scher K, Narducci S
A Protocol-Driven, Bedside Digital Conversational Agent to Support Nurse Teams and Mitigate Risks of Hospitalization in Older Adults: Case Control Pre-Post Study
J Med Internet Res 2019;21(10):e13440
URL: http://www.jmir.org/2019/10/e13440/
doi:10.2196/13440
PMID:31625949

Corrigenda and Addenda

Correction: Adherence and Satisfaction of Smartphone- and Smartwatch-Based Remote Active Testing and Passive Monitoring in People With Multiple Sclerosis: Nonrandomized Interventional Feasibility Study

Luciana Midaglia[1,2], MD; Patricia Mulero[1], MD; Xavier Montalban[1,3], MD, PhD; Jennifer Graves[4], MAS, MD, PhD; Stephen L Hauser[5], MD; Laura Julian[6], PhD; Michael Baker[7], MSc; Jan Schadrack[7], MD; Christian Gossens[7], MBA, PhD; Alf Scotland[7], MSc; Florian Lipsmeier[7], PhD; Johan van Beek[7], PhD; Corrado Bernasconi[7], MD, PhD; Shibeshih Belachew[7], MD, PhD; Michael Lindemann[7,8], MBA, PhD

[1]Department of Neurology-Neuroimmunology, Multiple Sclerosis Centre of Catalonia, Vall d'Hebron University Hospital, Barcelona, Spain
[2]Department of Medicine, Autonomous University of Barcelona, Barcelona, Spain
[3]Division of Neurology, University of Toronto, Toronto, ON, Canada
[4]Department of Neurology, University of California, San Diego, San Diego, CA, United States
[5]Department of Neurology, University of California, San Francisco, San Francisco, CA, United States
[6]Genentech Inc, South San Francisco, CA, United States
[7]F Hoffmann–La Roche Ltd, Basel, Switzerland
[8]Department of Economics, Baden-Wuerttemberg Cooperative State University, Loerrach, Germany

Corresponding Author:
Christian Gossens, MBA, PhD
F Hoffmann–La Roche Ltd
124 Grenzacherstrasse
Basel,
Switzerland
Phone: 41 61 687 5113
Fax: 41 61 691 9391
Email: christian.gossens@roche.com

Related Article:

Correction of: https://www.jmir.org/2019/8/e14863/

(J Med Internet Res 2019;21(10):e16287) doi:10.2196/16287

In "Adherence and Satisfaction of Smartphone- and Smartwatch-Based Remote Active Testing and Passive Monitoring in People With Multiple Sclerosis: Nonrandomized Interventional Feasibility Study" by Midaglia et al (J Med Internet Res 2019;21(8):e14863), there were inconsistencies in the phrases used to describe people with multiple sclerosis. These inconsistencies were inadvertently included during the production process.

The following revisions to the manuscript have been made:

- Throughout the manuscript, all instances of the phrase "persons with multiple sclerosis" have been revised to "people with multiple sclerosis", including one misspelled instance in the Abstract.
- In the first paragraph of the Introduction, the phrase "people with MS" has been revised to "people with multiple sclerosis".
- In Table 1, two instances of the acronym "PwMS" within the table have been changed to "people with multiple sclerosis". Furthermore, the table footnote which explained the acronym has been deleted: "[c]PwMS: people with multiple sclerosis."
- In Table 2, three instances of the acronym "PwMS" within the table have been changed to "people with multiple sclerosis". Furthermore, the table caption has been revised from "Demographics and characteristics of people with multiple sclerosis (PwMS) and healthy controls (HC) at baseline" to "Demographics and characteristics of people with multiple sclerosis and healthy controls (HCs) at baseline."
- The abbreviation list has been updated to remove the term "PwMS: people with multiple sclerosis".

Furthermore, all instances of "healthy controls" have been correctly abbreviated to "HCs" (whereas the instances in Table

2 and the definition in the Abbreviations list were previously missing the final “s”).

Lastly, the authors wish to correct an error in Table 2 which previously read “relaxing-remitting multiple sclerosis” but has now been changed to “relapsing-remitting multiple sclerosis”.

The changes made do not affect the findings of the study.

The correction will appear in the online version of the paper on the JMIR website on October 8, 2019, together with the publication of this correction notice. Because this was made after submission to PubMed, PubMed Central, and other full-text repositories, the corrected article has also been resubmitted to those repositories.

Submitted 17.09.19; this is a non–peer-reviewed article;accepted 23.09.19; published 08.10.19.

Please cite as:
Midaglia L, Mulero P, Montalban X, Graves J, Hauser SL, Julian L, Baker M, Schadrack J, Gossens C, Scotland A, Lipsmeier F, van Beek J, Bernasconi C, Belachew S, Lindemann M
Correction: Adherence and Satisfaction of Smartphone- and Smartwatch-Based Remote Active Testing and Passive Monitoring in People With Multiple Sclerosis: Nonrandomized Interventional Feasibility Study
J Med Internet Res 2019;21(10):e16287
URL: https://www.jmir.org/2019/10/e16287
doi:10.2196/16287
PMID:31596725

Corrigenda and Addenda

Addendum to the Acknowledgements: Mood Prediction of Patients With Mood Disorders by Machine Learning Using Passive Digital Phenotypes Based on the Circadian Rhythm: Prospective Observational Cohort Study

Chul-Hyun Cho[1*], MD, PhD; Taek Lee[2*], PhD; Min-Gwan Kim[3], MS; Hoh Peter In[3], PhD; Leen Kim[1], MD, PhD; Heon-Jeong Lee[1], MD, PhD

[1]Korea University College of Medicine, Department of Psychiatry, Seoul, Republic of Korea

[2]Sungshin University, Department of Convergence Security Engineering, Seoul, Republic of Korea

[3]Korea University College of Informatics, Department of Computer Science and Engineering, Seoul, Republic of Korea

[*]these authors contributed equally

Corresponding Author:
Heon-Jeong Lee, MD, PhD
Korea University College of Medicine
Department of Psychiatry
Anam Hospital, Korea University College of Medicine
Seoul
Republic of Korea
Phone: 82 029205815
Email: leehjeong@korea.ac.kr

Related Article:

Correction of: https://www.jmir.org/2019/4/e11029/

(J Med Internet Res 2019;21(10):e15966) doi:10.2196/15966

KEYWORDS

mood disorder; circadian rhythm; prediction; machine learning; digital phenotype; wearable device

The authors of "Mood Prediction of Patients With Mood Disorders by Machine Learning Using Passive Digital Phenotypes Based on the Circadian Rhythm: Prospective Observational Cohort Study" (J Med Internet Res 2019;21(4):e11029) missed an important source of funding in the Acknowledgments section — the Korea Health 21 R&D Project funded by the National Research Foundation of Korea (2017M3A9F1031220).

At the time of publication, the Acknowlegments section read as follows:

> *This independent research study was supported by the Korea Health 21 R&D Project funded by the Ministry of Health & Welfare, Republic of Korea (HM14C2606 and HI14C3212) and the National Research Foundation of Korea (2016M3C7A1904345). Most of all, the authors express their gratitude to the participants who took part in the study.*

The revised Acknowledgments section now appears as follows:

> *This independent research study was supported by the Korea Health 21 R&D Project funded by the Ministry of Health & Welfare, Republic of Korea (HM14C2606 and HI14C3212), and the National Research Foundation of Korea (2016M3C7A1904345 and 2017M3A9F1031220). Most of all, we express our gratitude to the participants who took part in the study.*

The correction will appear in the online version of the paper on the JMIR website on October 3, 2019, together with the publication of this correction notice. Because this was made after submission to PubMed, PubMed Central, and other full-text repositories, the corrected article has also been resubmitted to those repositories.

Submitted 22.08.19; this is a non–peer-reviewed article;accepted 29.08.19; published 03.10.19.

Please cite as:
Cho CH, Lee T, Kim MG, In HP, Kim L, Lee HJ
Addendum to the Acknowledgements: Mood Prediction of Patients With Mood Disorders by Machine Learning Using Passive Digital Phenotypes Based on the Circadian Rhythm: Prospective Observational Cohort Study
J Med Internet Res 2019;21(10):e15966
URL: https://www.jmir.org/2019/10/e15966
doi:10.2196/15966
PMID:31584007

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org