Journal of Medical Internet Research

Journal Impact Factor (JIF) (2023): 5.8
Volume 19 (2017), Issue 9   ISSN 1438-8871   Editor in Chief: Gunther Eysenbach, MD, MPH

Contents

Review

Components and Outcomes of Internet-Based Interventions for Caregivers of Older Adults: Systematic Review (e313)
Cassioppée Guay, Claudine Auger, Louise Demers, W Mortenson, William Miller, Dominique Gélinas-Bronsard, Sara Ahmed. 4

Viewpoint

A Dermatologist's Ammunition in the War Against Smoking: A Photoaging App (e326)
Titus Brinker, Alexander Enk, Martina Gatzka, Yasuhiro Nakamura, Wiebke Sondermann, Albert Omlor, Maximilian Petri, Ante Karoglan, Werner Seeger, Joachim Klode, Christof von Kalle, Dirk Schadendorf. 22

Original Papers

Effectiveness of a Web-Based Tailored Intervention With Virtual Assistants Promoting the Acceptability of HPV Vaccination Among Mothers of Invited Girls: Randomized Controlled Trial (e312)
Mirjam Pot, Theo Paulussen, Robert Ruiter, Iris Eekhout, Hester de Melker, Maxine Spoelstra, Hilde van Keulen. 26

User Acceptance of Computerized Cognitive Behavioral Therapy for Depression: Systematic Review (e309)
Theresia Rost, Janine Stein, Margrit Löbner, Anette Kersting, Claudia Luck-Sikorski, Steffi Riedel-Heller. 44

Personalized Hypertension Management Using Patient-Generated Health Data Integrated With Electronic Health Records (EMPOWER-H): Six-Month Pre-Post Study (e311)
Nan Lv, Lan Xiao, Martha Simmons, Lisa Rosas, Albert Chan, Martin Entwistle. 57

Web-Based Aftercare for Women With Bulimia Nervosa Following Inpatient Treatment: Randomized Controlled Efficacy Trial (e321)
Corinna Jacobi, Ina Beintner, Eike Fittig, Mickey Trockel, Karsten Braks, Carmen Schade-Brittinger, Astrid Dempfle. 73

Photoaging Mobile Apps in School-Based Melanoma Prevention: Pilot Study (e319)
Titus Brinker, Christian Brieske, Christoph Schaefer, Fabian Buslaff, Martina Gatzka, Maximilian Petri, Wiebke Sondermann, Dirk Schadendorf, Ingo Stoffels, Joachim Klode. 88

Prescribing of Electronic Activity Monitors in Cardiometabolic Diseases: Qualitative Interview-Based Study (e328)
Alice Bellicha, Sandrine Macé, Jean-Michel Oppert. 98

Social Media for the Dissemination of Cochrane Child Health Evidence: Evaluation Study (e308)
Michele Dyson, Amanda Newton, Kassi Shave, Robin Featherstone, Denise Thomson, Aireen Wingert, Ricardo Fernandes, Lisa Hartling. . . 106

Cardiac Patients' Experiences and Perceptions of Social Media: Mixed-Methods Study (e323)
Stephanie Partridge, Anne Grunseit, Patrick Gallagher, Becky Freeman, Blythe O'Hara, Lis Neubeck, Sarah Due, Glenn Paull, Ding Ding, Adrian Bauman, Philayrath Phongsavan, Kellie Roach, Leonie Sadler, Helen Glinatsis, Robyn Gallagher. . . . 118

Blending Face-to-Face and Internet-Based Interventions for the Treatment of Mental Disorders in Adults: Systematic Review (e306)
Doris Erbe, Hans-Christoph Eichert, Heleen Riper, David Ebert. . . . 131

Enhancing Seasonal Influenza Surveillance: Topic Analysis of Widely Used Medicinal Drugs Using Twitter Data (e315)
Ireneus Kagashe, Zhijun Yan, Imran Suheryani. . . . 144

The User Knows What to Call It: Incorporating Patient Voice Through User-Contributed Tags on a Participatory Platform About Health Management (e292)
Annie Chen, Rachel Carriere, Samantha Kaplan. . . . 158

Evaluating Patient Empowerment in Association With eHealth Technology: Scoping Review (e329)
Tracie Risling, Juan Martinez, Jeremy Young, Nancy Thorp-Froslie. . . . 174

Effectiveness of Telemonitoring in Obstetrics: Scoping Review (e327)
Dorien Lanssens, Thijs Vandenberk, Inge Thijs, Lars Grieten, Wilfried Gyselaers. . . . 184

Barriers and Facilitators to eHealth Use in Daily Practice: Perspectives of Patients and Professionals in Dermatology (e300)
Lieneke Ariens, Florine Schussler-Raymakers, Cynthia Frima, Annebeth Flinterman, Eefje Hamminga, Bernd Arents, Carla Bruijnzeel-Koomen, Marjolein de Bruin-Weller, Harmieke van Os-Medendorp. . . . 196

The Digital Divide and Health Disparities in China: Evidence From a National Survey and Policy Implications (e317)
Y Hong, Zi Zhou, Ya Fang, Leiyu Shi. . . . 206

Internet Searches and Their Relationship to Cognitive Function in Older Adults: Cross-Sectional Analysis (e307)
Johanna Austin, Kristy Hollingshead, Jeffrey Kaye. . . . 219

Self-Monitoring Kidney Function Post Transplantation: Reliability of Patient-Reported Data (e316)
Céline van Lint, Wenxin Wang, Sandra van Dijk, Willem-Paul Brinkman, Ton Rövekamp, Mark Neerincx, Ton Rabelink, Paul van der Boog. . . 230

Putting the Focus Back on the Patient: How Privacy Concerns Affect Personal Health Information Sharing Intentions (e169)
Mohamed Abdelhamid, Joana Gaia, G Sanders. . . . 245

Reliability of an e-PRO Tool of EORTC QLQ-C30 for Measurement of Health-Related Quality of Life in Patients With Breast Cancer: Prospective Randomized Trial (e322)
Markus Wallwiener, Lina Matthies, Elisabeth Simoes, Lucia Keilmann, Andreas Hartkopf, Alexander Sokolov, Christina Walter, Nina Sickenberger, Stephanie Wallwiener, Manuel Feisst, Paul Gass, Peter Fasching, Michael Lux, Diethelm Wallwiener, Florin-Andrei Taran, Joachim Rom, Andreas Schneeweiss, Joachim Graf, Sara Brucker. . . . 257

Computerized Adaptive Testing Provides Reliable and Efficient Depression Measurement Using the CES-D Scale (e302)

Bao Loe, David Stillwell, Chris Gibbons. 272

How the “Understanding Research Evidence” Web-Based Video Series From the National Collaborating Centre for Methods and Tools Contributes to Public Health Capacity to Practice Evidence-Informed Decision Making: Mixed-Methods Evaluation (e286)

Linda Chan, Jeannie Mackintosh, Maureen Dobbins. 285

Video Game Intervention for Sexual Risk Reduction in Minority Adolescents: Randomized Controlled Trial (e314)

Lynn Fiellin, Kimberly Hieftje, Tyra Pendergrass, Tassos Kyriakides, Lindsay Duncan, James Dziura, Benjamin Sawyer, Linda Mayes, Cindy Crusto, Brian Forsyth, David Fiellin. 298

Review

Components and Outcomes of Internet-Based Interventions for Caregivers of Older Adults: Systematic Review

Cassioppée Guay[1,2], MScOT; Claudine Auger[1,2], PhD; Louise Demers[1,3], PhD; W Ben Mortenson[4,5,6], PhD; William C Miller[4,5], PhD; Dominique Gélinas-Bronsard[1,2], BScOT; Sara Ahmed[2,7], PhD

[1]School of Rehabilitation, Faculty of Medicine, Université de Montréal, Montreal, QC, Canada

[2]Centre for Interdisciplinary Research in Rehabilitation of Greater Montreal, Montreal, QC, Canada

[3]Centre de recherche de l'Institut universitaire de gériatrie de Montréal, Montreal, QC, Canada

[4]Department of Occupational Science and Occupational Therapy, Faculty of Medicine, University of British Columbia, Vancouver, BC, Canada

[5]GF Strong Rehabilitation Center, Vancouver, BC, Canada

[6]International Collaboration on Repair Discoveries, Vancouver, BC, Canada

[7]School of Physical and Occupational Therapy, Faculty of Medicine, McGill University Health Centre, Clinical Epidemiology, McGill University, Montreal, QC, Canada

Corresponding Author:
Cassioppée Guay, MScOT
School of Rehabilitation
Faculty of Medicine
Université de Montréal
7077
Avenue du Parc
Montreal, QC,
Canada
Phone: 1 514 796 4551
Email: cassioppee.guay@gmail.com

Abstract

Background: When trying to access interventions to improve their well-being and quality of life, family caregivers face many challenges. Internet-based interventions provide new and accessible opportunities to remotely support them and can contribute to reducing their burden. However, little is known about the link existing between the components, the use of behavior change techniques, and the outcomes of these Internet-based interventions.

Objective: This study aimed to provide an update on the best available evidence about the efficacy of Internet-based interventions for caregivers of older adults. Specifically, the components and the use of behavior change techniques and how they impact on the efficacy of the intervention were sought.

Methods: A systematic review searched primary source studies published between 2000 and 2015. Included studies were scored with a high level of evidence by independent raters using the GRADE criteria and reported caregiver-specific outcomes about interventions delivered through the Internet for caregivers of people aged 50 years and older. A narrative synthesis identified intervention components (eg, content, multimedia use, interactive online activities, and provision of support), behavior change techniques, and caregiver outcomes (eg, effects on stressors, mediators, and psychological health). The risk of bias within the included studies was assessed.

Results: A total of 2338 articles were screened and 12 studies describing 10 Internet-based interventions were identified. Seven of these interventions led to statistically significant improvements in caregiver outcomes (eg, reducing depression or anxiety, n=4). These efficacious interventions used interactive components, such as online exercises and homework (n=4) or questionnaires on health status (n=2) and five of them incorporated remote human support, either by professionals or peers. The most frequently used behavior change techniques included in efficacious interventions were provision of social support (n=6) and combinations of instructions to guide behavior change and barrier identification (n=5). The design and aim of the included studies did not permit determining exactly which component and/or behavior change technique was more efficacious in producing positive outcomes in caregivers. The risk for selection bias was low for all the studies, and low to high for performance, detection, and attrition biases.

Conclusions: In sum, Internet-based interventions that incorporate professional and social support, and provide instructions to change behavior and problem solve in an interactive manner appear to lead to positive outcomes in caregivers. Studies isolating the specific effect of components are needed to improve our understanding of the underlying mechanism of action.

(J Med Internet Res 2017;19(9):e313) doi:10.2196/jmir.7896

KEYWORDS

systematic review; caregivers; aged; Internet-based interventions; Internet; behavior change

Introduction

A family caregiver (henceforth described as caregivers) is a person who provides care without any financial compensation for a family member, a friend, or a loved one with long-term health problems or disabilities [1]. Approximately 5.4 million Canadians [2] and 34.2 million Americans [3] are caregivers of older adults. Aging and associated frailty are the main causes of caregiving needs [2,3], which are expected to increase because of population aging in Western countries. Moreover, older adults are more prone to experience disability related to long-term physical conditions and memory problems [3,4], which altogether increase the caregiver burden. The term “burden” is used to describe the harmful physical, emotional, and social effects caregivers may experience [5]. Role overload and burden put caregivers at higher risk for depression, anxiety, and negative levels of stress [6,7]. In 2012 in Canada, one in two caregivers of an older adult felt anxious about their responsibilities and one in six experienced depressive symptoms [8]. In the United States, 17% of the caregivers of older adults report poor or fair health and one in five report a decrease in their health due to caregiving [3].

Stress Process Model

The Stress Process Model [9] is a conceptual framework that is used to explicate the stress and burden experienced by caregivers. Caregiving may produce primary stressors, which are situations that are perceived as problematic or harmful by the caregiver. These stressors can be described objectively (eg, low level of independence of the care recipient, observable behaviors) and subjectively (eg, feelings of overload and relational deprivation). If stressors persist, they may lead to secondary role strains (eg, job-caregiving conflicts, constriction of social life) and intrapsychic strains (eg, decreased self-esteem and sense of mastery, loss of self). Primary stressors and secondary strains contribute to overall caregiver burden, which can result in adverse outcomes on psychological and physical health as well as on social participation. According to the Stress Process Model, personal resources, such as coping strategies and access to social support, are mediators of the stress process that can help mitigate adverse outcomes. The model also acknowledges that caregiving can lead to positive effects, such as a sense of inner growth while facing challenges.

Considering the stress and burden they experience, caregivers of older adults require health care, psychosocial, community, and respite services to prevent negative outcomes related to their caregiving role. However, caregivers experience many barriers when trying to obtain those services, such as lack of transportation to access the intervention, unavailability of a secondary caregiver to take over in their absence, and lack of flexibility to participate in a highly demanding intervention [10]. Consequently, programs that are delivered outside of the home setting have been shown to be less attended to by caregivers than home-based programs, such as telephone counseling, home visits, and technology-based interventions [10]. Internet-based interventions can thus offer an easily accessible alternative [11-17] and can be more cost-effective than traditional face-to-face interventions [18].

Internet-Based Interventions for Caregivers

Internet-based interventions, also referred to as eHealth interventions or information and communication technology-based interventions, are defined as therapeutic programs with specific health objectives delivered mainly using the Internet [19]. They have been classified by Barack and Klein [19] into six categories: Web-based education interventions, self-help Web-based therapeutic interventions, human-supported Web-based therapeutic interventions, online counseling, Internet-operated therapeutic software, and other online activities. Each category of Internet-based intervention is described with respect to four major types of components: (1) content (eg, educational or aimed for behavior change), (2) multimedia (eg, text, images, videos), (3) interactive online activities (eg, online quizzes, homework), and (4) guidance and supportive feedback (eg, automatic reminders, professional feedback).

To our knowledge, seven reviews [11-17] have described the efficacy of Internet-based interventions for caregivers of older adults, but the portrait they provide is incomplete regarding the quality of the evidence and the components driving the success of the interventions. The reviews included Internet-based interventions for caregivers of all ages [12,13] or caring for people with cancer [14], dementia [11,17], for community-dwelling older adults [15], or for adults and older adults with intellectual disabilities [16]. The reviews report that Internet-based interventions can effectively reduce depression and caregiver burden [11-14,16,17], as well as having positive effects on caregivers’ sense of competence/self-efficacy [11,17], coping strategies [14], knowledge about the care recipient’s condition [17], and quality of life [14]. Lastly, the review by Magnusson et al [15] gave interesting insights on success factors and feasibility of Internet-based interventions, universal design principles, and older adults’ thoughts and attitudes toward technology, but not on the efficacy of included interventions. Concerning methodological aspects, only two reviews produced a complete analysis of the quality of evidence, which was deemed poor [11] or acceptable [13]. Other reviews either described the quality of the evidence using scales [12,14], but did not use this assessment to critically appraise the strength of reported results or did not report anything on the study designs or risk of bias [15-17]. None of the seven reviews used a framework to classify the components, thus making it difficult to compare components of efficacious Internet-based interventions and to identify hypothetical causes of efficacy. In sum, the currently available reviews indicate that Internet-based interventions for caregivers of older adults facing different health conditions have the potential to generate positive outcomes on psychological health. The strength of evidence of these results, however, is uncertain and reasons of observed improvements in caregivers' outcomes were poorly documented and not reported uniformly.

Behavior Change Techniques in Internet-Based Interventions

One factor that might explain the efficacy of Internet-based interventions, aside from the components of the intervention itself, is the incorporation of behavior change techniques (BCTs). BCTs are strategies that promote behavior change by, for example, providing information on consequences of behavior on health, prompting users to identify barriers to behavior change or offering social support. A review by Webb et al [20] found that Internet-based interventions promoting healthy living habits that are theoretically grounded and using BCTs are associated with larger effect sizes when compared with another intervention or no intervention at all. Similarly, a rapid review on the potential of Internet-based interventions for self-management argues that the incorporation of cognitive-behavioral therapy as well as BCTs is required to attain certain behavioral outcomes (eg, healthier living habits, safer sex) [21]. Given the encouraging results of integrating BCTs in the fields of health promotion and prevention for care recipients, the investigation whether Internet-based interventions for caregivers should enhance self-management by encouraging the development of new behaviors with the use of BCTs is relevant. For this, the taxonomy by Abraham and Michie [22], which proposed a series of BCTs to incorporate into interventions aiming to support behavior change, can be used to identify and classify BCTs as a way of pinpointing underlying mechanisms of effects. This has not been done in any previous reviews for caregivers of older adults to our knowledge.

Aim and Specific Objectives

This study aimed to provide an update on the best available evidence about the efficacy of Internet-based interventions for caregivers of older adults. Specific objectives were to (1) classify the components that are found in Internet-based interventions for caregivers of older adults, (2) describe the BCTs used in these Internet-based interventions, and (3) explore which intervention components and BCTs of Internet-based interventions are associated with efficacious outcomes in caregivers.

Methods

Information Sources and Search Method

Following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement [23], a systematic literature search was conducted in MEDLINE, EMBASE, and CINHAL, covering studies published from January 2000 to July 2015. With the help of a biomedical research librarian, a list of more than 50 medical subject headings, descriptors, and keywords, matched to the specific thesaurus of every database consulted, was used to identify the population (caregivers) and the interventions (Internet-based) of interest (see Multimedia Appendix 1). In addition, we used keywords related to the concepts of self-management and behavior change to target interventions with BCTs for achieving outcomes on the health and well-being of caregivers.

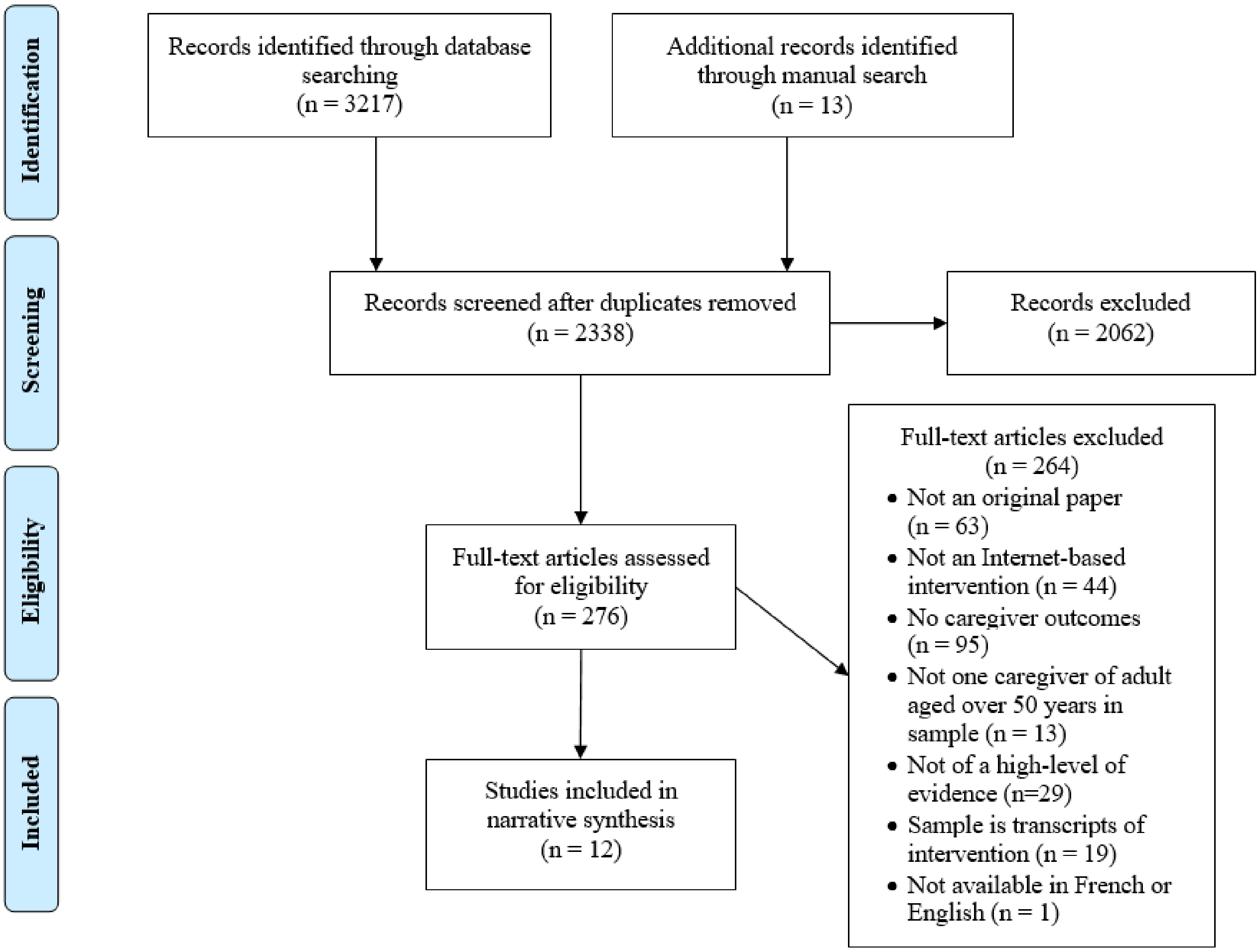
Eligibility Criteria

Studies were included if they (1) were original papers published in peer-reviewed journals, (2) reported on an intervention of which the principal mode of delivery was the Internet, (3) reported caregiver-specific outcomes, (4) targeted caregivers of older adults and thus had a sample including at least one caregiver of a person aged 50 years or older, and (5) were of a high level of evidence according to the Grading of Recommendations Assessment, Development and Evaluation (GRADE) criteria [24]. As per these criteria, the initial level of evidence was determined based on study designs (eg, high level for randomized controlled trials, moderate level for quasi-randomized controlled trials, and low levels for observational studies). Studies were scored lower if a combination of serious limitations were present (eg, lack of allocation concealment, lack of blinding, incomplete accounting of patients and outcomes events, selective reporting, and other limitations). Studies were scored higher than their original level if they displayed strong methodological qualities (eg, no plausible cofounders, no major treats to the validity of the results), if a dose-response gradient was found, or if the effect would have been reduced by cofounders. Studies were excluded if they were not written in either French or English. Descriptive studies analyzing data consisting of Internet transcripts (eg, messages on a forum, blogs, posting) were also excluded.

Study Selection

The complete selection process is detailed in Figure 1. Titles and abstracts were reviewed by the first author. Full-text articles were obtained for relevant studies and the preceding criteria were applied. When the first author was uncertain about the inclusion of studies, they were reviewed independently by a second author (DGB). If consensus could not be reached, discrepancies were resolved with the help of a third author (CA). A manual search was conducted in the reference lists of reviews found during this search.

Figure 1. PRISMA flowchart of the search strategy and results.



Data Collection and Coding

The first author abstracted data from included studies using Excel forms to record the study characteristics, components of the Internet-based interventions, and the use of BCTs. Caregiver population, experimental and control conditions, data collection for reported outcomes, analyses performed, and additional characteristics specific to Internet-based trials as outlined in the CONSORT-EHEALTH (Consolidated Standards of Reporting Trials) guidelines (eg, computer literacy, intended dosage, usage outcomes) [25] were documented.

Caregiver outcomes were classified with the Stress Process Model [9] to outline which factors of the stress process were measured: primary stressors, secondary role strains, secondary intrapsychic strain, or outcomes.

All Internet-based interventions and their components were coded using the Barack and Klein categorization [19]. Following this classification, interventions were classified in one of six possible categories. First were Web-based education interventions, which are online programs or websites providing general standardized content to improve knowledge, awareness, and understanding of users, but not their behavior. Second and third were self-help Web-based therapeutic interventions and human-supported Web-based therapeutic interventions, which are both multicomponent online interventions with content tailored to support behavior change. Those two distinguish themselves by the amount of guidance and feedback they provide: self-help interventions are designed to be self-guided by the user and so do not require interactions with humans, whereas human-supported interventions provide a variety of support means from professionals or peers. Online counseling is the fourth category and refers to programs or technologies that primarily enable professional-user communication for remote counseling, usually on subjects such as mental health or psychological follow-ups. Within the fifth category were Internet-operated therapeutic software, which are programs or devices that use advanced computer capabilities (eg, artificial intelligence principles, augmented realities, or algorithms) to produce a robotic therapist simulation providing dialog-based therapy with patients, rule-based systems/games (eg, WII console), or three-dimensional environments (eg, Second Life). Finally, the category of other online activities consisted of any other websites, blogs, informal support groups, wikis, podcasts, and self-assessments available on the Web that did not have a specified therapeutic goal. Components found from the description of interventions provided by the authors in the article were classified within the four major component categories, as described in Textbox 1.

Textbox 1. Components of Internet-based interventions (adapted from Barack and Klein [19]).

Content: nature of the information disseminated through the program. Can be generic and educative or designed to create a therapeutic change.

Multimedia: means used to disseminate the content (eg, text, graphics, video).

Interactive online activities: opportunities given to participate actively within the program (eg, quizzes, exercises, questionnaires).

Guidance and supportive feedback: tools by which users access external information about their performance and progress. Can be offered automatically with integrated algorithms (eg, reminders) or by professionals and/or peers through asynchronous (eg, email, forums, bulletin boards) or synchronous (eg, videoconference) components.

Behavior change techniques were also extracted from the description of interventions found in the report using the taxonomy of Abraham and Michie [22]. This taxonomy details 26 BCTs that can be incorporated in interventions, ranging from the simple provision of information on consequences of one's behavior to more complex techniques such as providing complete goal setting, modeling appropriate behavior, or providing detailed feedback on behavior performance.

Finally, although a high level of evidence per the GRADE criteria was a condition for inclusion, risk of bias was still assessed as “high,” “low,” or “unclear” for the random sequence generation and concealment of allocation (selection bias), for blinding of outcome assessors and the use of valid measures (detection bias), for the blinding of the participants (performance bias), and for how withdrawals were statistically accounted for (attrition bias). This was done as recommended by the Cochrane Handbook for Systematic Reviews of Interventions [26]. Given the difficult nature of double blinding in psychosocial and self-management intervention trials, which comprise the majority of Internet-based interventions for caregivers [12], the double-blind criterion was not assessed.

Results

After applying the search strategy detailed in Figure 1, a total of 12 studies were retained for a narrative synthesis [27-38]. Records were excluded mainly because they were not a primary source study, they were not delivered through the Internet, or they did not report caregiver-specific outcomes. Twenty-nine studies were excluded because of the quality of their study design; per the GRADE criteria, 8 were of a moderate level of evidence [39-46], 18 of a low level [47-64], and 3 were of a very low level [65-67]. Three reports were downgraded from a high level of evidence to a moderate level because of high differential rates of dropouts between control and experimental conditions that were not accounted for with proper intention-to-treat analysis [40,42,43]. None of the studies were upgraded.

Characteristics of the Included Studies

The main characteristics of the included studies are presented in Tables 1 and 2. Regarding the studied population, caregivers were mainly female adults or older adults and generally were either a spouse or a child of the care recipient. They cared for people with dementia [27,28,30,38], stroke [32,33,36,37], cancer [29,31,35], or traumatic brain injury [34].

Table 1. Characteristics of the study population of included studies (N=12).

Author	Country	N	Mean age (SD)[b]	Female[b]	Relationship with care recipient[b]	Diagnosis of care recipient[b]
Beauchamp et al [27]	USA	299	46.9 (12.2)	73%	Child (67%)	Dementia
Blom et al [28]	Netherlands	245	61.2 (12.37)	69.4%	Spouse (58.4%)	Dementia
Chih et al [29]	USA	235	56	64.2%	Spouse/partner (69.3%)	Cancer
Cristancho-Lacroix et al [30]	France	49	64.2 (10.3)[a]	16 (64%)[a]	Child (64%)[a]	Dementia
DuBenske et al [31]	USA	246	55.56	68.3%	Spouse/partner (72%)	Cancer
Eames et al [32]	Australia	61	55.5	64%	Spouse/partner (67%)	Stroke
Kim et al [33]	South Korea	36	53 (13.7)	NR	Spouse (66.7%)[a]	Stroke
McLaughlin et al [34]	USA	201	NR, 34.6% aged 51-60[a]	86.4%[a]	NR	TBI
Namkoong et al [35]	USA	285	55.56	68.3%	NR	Cancer
Pierce et al [36]	USA	103	54 (12.2)[a]	69.4%[a]	Wife (41.7%)[a]	Stroke
Smith et al [37]	USA	32	55.3 (6.9)[a]	100%	Wife (100%)	Stroke
Torkamani et al [38]	UK, Spain, and Greece	60	60.69 (13.90)	45%	NR	Dementia

[a] For intervention group only.

[b] NR: none reported; SD: standard deviation; TBI: traumatic brain injury.

Table 2. Description of the intervention and control groups of included studies (N=12).

Author	Intervention			Control group	
	n[a]	Duration[b]	Description	n[a]	Description
Beauchamp et al [27]	150	30 days	Caregiver's Friend: Dealing with Dementia—an ongoing worksite Web-based support program providing materials tailored to the needs of caregivers in 3 distinct modules (being a caregiver, coping with emotions, and common difficulties)	149	Usual care wait list
Blom et al [28]	149	5-6 months	Mastery Over Dementia: a 9-lesson online program; the first 8 lessons followed the same sequence: provision of information, exercises, homework, and feedback; lessons were about coping with behavioral problems, relaxation, arranging help from others, changing nonhelping thoughts, and communication; final lesson was a recap and booster session	96	E-bulletin sent by email every 3 weeks for 6 months; content did not overlap with intervention
Chih et al [29]	118	12-24 months	Comprehensive Health Enhancement Support System (CHESS): a password-protected website in which users self-directed to a variety of services (information, communication, and coaching); content covered cancer, caregiving and palliative care, emotional distress, use of coping techniques, and communication techniques	117	Access to the same intervention, without one component (clinical report) for 12-24 months
Cristancho-Lacroix et al [30]	25	12 weeks[c]	Diapason: a password-protected website offering information, skills training and a forum for caregivers; content was divided in 12 thematic sessions with videos covering caregiver stress, understanding the disease, maintaining the loved ones' autonomy, understanding their reactions, coping with behavioral and emotional troubles, communicating, improving their daily lives, avoiding falls, pharmacological and nonpharmacological interventions for caregivers, social and financial support, and about the future	24	Usual care
DuBenske et al [31]	124	2 years or up to 13 months after the death of the care recipient[c]	Comprehensive Health Enhancement Support System (CHESS): a password-protected website in which users self-directed to a variety of services (information, communication, and coaching); content covered cancer, caregiving and palliative care, emotional distress, use of coping techniques, and communication techniques	122	Access to a list of cancer and palliative care websites constructed from the opinions of clinicians in addition to usual care
Eames et al [32]	31	3 months[c]	What You Need to Know About Stroke: an educational package online containing a list of 34 topics regarding stroke; the Web-based intervention was reinforced with 3 face-to-face and 3 telephone meetings with participants	30	Usual care for the care recipient
Kim et al [33]	18	9 weeks	A Web-based program incorporating education and resources to support self-efficacy in the home setting. Content was divided in nine video sessions covering three themes: understanding stroke, recurrence prevention, and family life	18	Access to an e-bulletin over the course of 6 months
McLaughlin et al [34]	104	3 months	Brain Injury Partner: a Web-based program designed to improve family advocacy skills with content covering advocacy skills, strategies for reducing stress, and to determine necessary professional support needs	97	Access to the Brain Injury Association of America (BIAUSA)
Namkoong et al [35]	141	2 years	Comprehensive Health Enhancement Support System (CHESS): a password-protected website in which users self-directed to a variety of services (information, communication, and coaching); content covered cancer, caregiving and palliative care, emotional distress, use of coping techniques, and communication techniques	144	Access to a list of high-quality patient-directed cancer and palliative care websites in addition to usual care
Pierce et al [36]	51	1 year	Caring-Web: an educational and support intervention that answered questions, discussed options, and gave up-to-date information covering frequently requested topics like stroke disease process, safe transfer techniques, and emotional changes	52	Specific instructions to not buy or use Internet during the study in addition to usual care

Author	Intervention			Control group	
	n[a]	Duration[b]	Description	n[a]	Description
Smith et al [37]	15	11 weeks	A Web-based conferencing and video education intervention designed to provide the caregiver with knowledge, resources, and skills; content was divided in 9 weekly video topics covering how to get in touch with your feelings as a caregiver, understanding what it's like to be a care recipient, being a good listener, nonverbal behavior, choice/control/predictability, relaxation and positive imagery to control stress, and the role of pleasant activities	17	Access only to one component of the intervention that presents links to resources
Torkamani et al [38]	30	6 months	A technology pLatform for the Assisted living of Dementia elDerly Individuals and their carers (ALADDIN): a Web-based program designed to provide support and information with content covering dementia and relaxation/exercises techniques	30	No attention or intervention given

[a] Before attrition.

[b] Length of access to intervention.

[c] In addition to usual care.

One Internet-based intervention was designed specifically to answer the needs of caregivers who were also workers [27]. None of the other studies mentioned the working status of caregivers. Three studies [29,31,35] assessed caregiver comfort levels with using the Internet on a five-point scale ranging from not comfortable to extremely comfortable. On the whole, caregivers rated themselves as being somewhat comfortable with using the Internet (mean scores ranging from 2.36 to 2.54). One study enrolled only novice users of the Internet [36].

The setting of the interventions varied across the included studies in terms of dosage, comparison conditions, and reported adherence. Half of the Internet-based interventions were administered to the experimental group without specification regarding dosage or a "use as you will" instruction [27,29,31,35,36,38], meaning that users could use the intervention whenever they wanted and how long they wanted. The other studies provided a set of explicit directions, such as requiring the user to log in at least once every week to view certain content [30,33] or for specific amounts of time [34,37]. Duration of interventions varied between 30 days [27] to 2 years [29,31,35]. Three control conditions were found within the included studies: (1) access to online resources (eg, e-bulletin on specific subjects, list of websites) [28,31,33-35], (2) access to selected portions of the experimental intervention [29,37], or (3) usual care/wait list [27,30,32,36,38]. Usual care for caregivers was either defined as the provision of information and education regarding care for the loved one [30,32] or not defined at all [31,33,36]. Adherence and usage were reported in six studies [27,30,31,33,34,36] and varied across interventions. For example, Kim et al [33] reported 100% adherence to requested usage (eg, all participants completed all the nine sessions planned in the program), whereas DuBenske et al [31] reported that 73% of the participants logged in at least once during the study.

As detailed in Table 3, the risk for selection bias was low for all the studies, and low to high for performance, detection, and attrition biases according to the tool of the Cochrane handbook [26]. Reasons for higher risk of bias concerned blinding of participants and outcome assessors, lack of control for co-intervention, and high rates of dropouts. Risk of bias in how participants performed and rated their health status at the time of completion was high overall, as only one study blinded their participants to group allocation [28] and one controlled for co-intervention [30]. A possibility for detection bias was also present for all studies because all assessments used for reporting the outcomes were based on self-report. This could have also led to a social desirability bias. Furthermore, the study of Beauchamp et al [27] did multiple testing without apparent statistical corrections, which can also lead to a detection bias. One study was judged to present a very high risk for detection bias because the outcome assessor was not blinded and the measures were administered during a face-to-face session [30]. As for the dropout rates, they were of 30% or more in both groups in five studies [29-31,35,36] and differed by 10% or more between the experimental and control groups in two studies [28,37]. Reasons for dropouts were always explored and missing data were treated with proper statistical analysis (eg, imputation techniques, intention-to-treat analysis, statistical models). Although these measures were taken, risk of bias for attrition was still judged high for these studies with large rates of dropouts and/or differential rates between experimental and control groups.

Table 3. Risk of bias.[a]

Author	Selection bias		Performance bias	Detection bias	Attrition bias
	Random sequence generation	Allocation concealment	Single blind	Blinding of outcome assessor	Missing data
Beauchamp et al [27]	+	+	–	+	+
Blom et al [28]	+	+	+	+	–
Chih et al [29]	+	+	–	+	–
Cristancho-Lacroix et al [30]	+	+	–	–	–
DuBenske et al [31]	+	+	–	+	–
Eames et al [32]	+	+	–	+	+
Kim et al [33]	+	+	–	?	+
McLaughlin et al [34]	+	+	–	+	+
Namkoong et al [35]	+	+	–	+	–
Pierce et al [36]	+	+	–	?	–
Smith et al [37]	+	+	–	+	–
Torkamani et al [38]	+	+	–	?	?

[a] +: low risk of bias; – high risk of bias; ?: unclear risk of bias.

Categories and Components of Internet-Based Interventions

The included studies reported results concerning 10 interventions because the outcomes of one intervention were reported in three different articles [29,31,35]. Three intervention categories were found: Web-based education interventions [32,36,38], self-help Web-based therapeutic interventions [27,34], and human-supported Web-based therapeutic interventions [28-31,33,35,37]. There were neither online counseling activities nor Internet-operated therapeutic software in our sample. The components, as categorized by Barack and Klein [19], are presented in Table 4 and concern the use of multimedia, interactive online activities, and the provision of guidance and support, either automatically by the program or by a human (peer or professional).

Multimedia

All interventions used written text as their main multimedia component and some also used videos [27,28,30,33,34,37]. Videos were skill-based or educational; in the intervention by McLaughlin et al [34], the videos were designed to teach the caregivers a specific set of advocacy skills, whereas in the intervention by Kim et al [33] the videos were recorded lectures, supported by PowerPoint presentations, intended for the caregiver to watch and learn about various topics related to stroke.

Interactive Online Activities

Most interventions offered interactive online activities, either in the form of homework, quizzes, and exercises to reinforce the educational content [28,33,34,37], or in the form of online questionnaires [27,29,31,35,38]. In the intervention by Beauchamp et al [27], online questionnaires asked about caregiver status to tailor the intervention so that participants would view only relevant content for their situation. For example, selecting the “spouse” status revealed content about finances, socializing, and losing a companion. Similarly, online questionnaires about the caregiver and care recipient’s health status in a multicomponent intervention for caregivers of people with cancer [29,31,35] were used for tailoring. In this intervention, answers were also compiled by the program and shared with the clinical team in the form of a clinical report, including graphics visually representing the answers and the evolution in the dyad’s health status. It was also the only Internet-based intervention with a decision aid system, called the “coaching service,” which offered caregivers detailed action plans and instructions to change behavior based on a detailed analysis of the responses to online questionnaires.

Guidance and Supportive Feedback

Human support was given asynchronously (eg, forum, email, bulletin boards) [28-31,33,35-38] and synchronously (eg, live chat session) [37]. Human support was offered by a health professional (eg, nurse, clinician, or psychologist) in six interventions [28,29,31,33,35-38] to address questions and problems caregivers might have during the intervention. The intervention by Blom et al [28] offered professional feedback regarding caregiver homework and exercises, and participants could progress in the intervention only if they opened and checked the feedback. Five interventions also offered opportunities for peer support [29-31,35-38], with the intervention by Smith et al [37] being the only one using a group-support context with live chat sessions guided by a professional.

Other common components included the provision of links to additional resources [29,31,33-37] and written or videotaped testimonials of other caregivers [27,29-31,35].

Table 4. Component categories[a] for each category of the Internet-based interventions.

Author	Multimedia	Interactive online activities[b]	Guidance and supportive feedback[b]	Other[b]
Web-based education interventions				
Eames et al [32]	Text	NR	NR	NR
Pierce et al [36]	Text	NR	Professional support: nurse specialist and rehabilitation team respond to questions with a private asynchronous module (email forum); peer support: asynchronous discussions facilitated by a nurse (email)	List of relevant Web links
Torkamani et al [38]	Text	Online questionnaires on CR and CG health status	Professional support: clinicians receive answers from IOA, facilitating the speedy delivery of appropriate interventions; clinicians are also reachable with a “contact us” button; peer support: asynchronous discussion sessions (forum)	Musical entertainment; relaxation and exercise techniques
Self-help Web-based therapeutic interventions				
Beauchamp et al [27]	Text; videos	Online questionnaires on CG personal situation; changing role button to select the relationship with CR	NR	IOA used to tailor content; testimonials
McLaughlin et al [34]	Text; videos	Video-based skills exercises	NR	List of relevant Web links and articles
Human-supported Web-based therapeutic interventions				
Blom et al [28]	Text; videos	Homework and exercises online; evaluation at the start and end of each lesson	Professional support: psychologist provides asynchronous feedback on IOA (electronic secured app); automatic reminders to send homework or attend lessons	Consultation of feedback is mandatory to have access to the next lesson
Cristancho-Lacroix et al [30]	Text; videos lectures	NR	Peer support: asynchronous discussion sessions moderated by a psychologist (forum)	Relaxation training; testimonials; glossary; bank of activities to stimulate CR
Chih et al [29]; DuBenske et al [31]; Namkoong et al [35]	Text; graphic	Online questionnaires on CR and CG health status; coaching service that automatically generates graphics of health status, offer decision aids, and structures an action plan	Professional support: cancer information specialist available via an “ask and expert” button.; Clinician report: summaries of users’ health available to the clinical team on demand, from a threshold alert or two days before a clinic visit; peer support: asynchronous discussion sessions moderated by a professional facilitator (bulletin board)	IOA and interactions through supportive feedback component used to tailor content; FAQs; list of relevant Web links, articles and community services; cancer news; testimonials ; personal webpage
Kim et al [33]	Video lectures; PowerPoint slides	Online quizzes following the viewing of video lectures	Professional support: asynchronous service to network with health professionals (email)	List of relevant Web links
Smith et al [37]	Text; video of enacted support group	At-home apps given by a nurse	Professional support: two times per week, a synchronous chat session directed by a nurse for the viewing and commenting of the weekly video (Adobe connect); the nurse is also available by asynchronous communication (email); peer support: asynchronous discussion sessions (email and message board)	List of relevant Web links, instructional videos and PDF files; online library of educational information; search engine

[a] As categorized by Barack and Klein [19].

[b] CG: caregiver; CR: care recipient; FAQ: frequently asked question: IOA: interactive online activities; NR: none reported.

Behavior Change Techniques

All three Web-based education interventions used less than two BCTs [32,36,38] (Table 5). All the other interventions incorporated four to 10 BCTs each. Overall, 15 BCTs were used within our sample, out of the 26 possibilities of Abraham and Michie's taxonomy [22]. The most commonly used technique was "social support" and this was offered through the component of peer support [29-31,35-38], except for one intervention that trained caregivers in planning social support instead of offering it to them [27]. Half of the interventions also provided a combination of two techniques: "providing instructions" and "prompting barrier identification" [27-31,34,35]. These instructions were given by multimedia components and/or reinforced by interactive online activities and professional support. "Stress management techniques" [28-31,34,35,37,38], "prompting practice of behavior" [30,34,37], and "modeling or demonstrating behavior" [27,28,34] were other common techniques used. The multicomponent intervention for caregivers of people with cancer [29,31,35] was the only intervention using "goal setting" and "action planning" techniques with their coaching service component. The intervention by Smith et al [37] incorporated the highest number of BCTs (n=10). Specifically, caregivers had to watch videos of enacted support group, which offered opportunities for social comparison and identification with role models. They also had to complete homework and report on their performance in subsequent discussion sessions to receive comments from professionals and peers, which prompted practice of behavior, self-monitoring, and provided feedback on performance.

Table 5. Behavior change techniques for each category of Internet-based interventions.

Author	Behavior change techniques[a]	Caregiver outcomes (ES)[b]
Web-based education interventions		
Eames et al [32]	NR	NSSD in caregiver strain
Pierce et al [36]	NR	NSSD in depression symptoms and satisfaction with life
Torkamani et al [38]	Social support; stress management	NSSD in caregiver burden, occurrence of psychiatric, and/or behavioral problems, depressive symptoms, and quality of life
Self-help Web-based therapeutic interventions		
Beauchamp et al [27]	Barrier identification; instructions; modeling; social support	↓ stress (0.5); ↑ intention to get support (0.3); ↓ caregiver strain (0.2); ↑ caregiver gain (0.2); ↓ depressive symptoms (0.2); ↓ state anxiety (0.2); ↑ self-efficacy (0.2); NSSD in the use of specific stress-reduction strategies
McLaughlin et al [34]	Barrier identification; instructions; modeling; prompt practice; stress management	↑ skill application (1.01); ↑ intention to use (0.7); ↑ knowledge (0.67); NSSD in satisfaction with life
Human-supported Web-based therapeutic interventions		
Blom et al [28]	Barrier identification; instructions; modeling; feedback on performance; stress management; time management	↓ symptoms of anxiety (0.48); ↓ depressive symptoms (0.26)
Cristancho-Lacroix et al [30]	Information on behavior-health link and on consequences; barrier identification; instructions; prompt practice; social comparison; social support; stress management	↑ knowledge (0.79); NSSD in perceived stress
Chih et al [29]; DuBenske et al [31]; Namkoong et al [35]	Information on behavior-health link and on consequences; barrier identification; instructions; goal setting; social support; stress management; time management	↓ negative mood at 6 and 12 months; ↓ caregiver burden at 6 months; ↑ bonding = ↑ active coping; NSSD for preparedness, physical burden, and in levels of disruptiveness
Kim et al [33]	Information on behavior-health link and on consequences; instructions; feedback on performance	↑ caregiver mastery
Smith et al [37]	Information on behavior-health link and on consequences; intention formation; instructions; self-monitoring of behavior; feedback on performance; prompt practice; social comparison; social support; identification to role models; stress management	↓ depression at 11 weeks and 1 month follow-up; NSSD in sense of mastery, self-esteem, and social support

[a] As categorized by Abraham and Michie [22].

[b] Arrows show the direction of statistically significant differences in intervention group compared to control for outcomes measured (*P*<.05). ES: value of effect sizes as originally reported by the authors; NR: none reported; NSSD: not statistically significant difference.

Table 6. Classification of the statistically significant outcomes by categories of Internet-based interventions and according to the Stress Process Model.[a]

Outcome	Web-based education interventions			Self-help Web-based therapeutic interventions		Human-supported Web-based therapeutic interventions				
	[32]	[36]	[38]	[27]	[34]	[28]	[30]	[29,31,35]	[33]	[37]
Primary stressors										
Problematic behavior							0			
Relationship quality							0[b]			
Secondary role strains										
Disruptiveness								0		
Secondary intrapsychic strains										
Caregiver gain				+						
Mastery								0	+	0
Self-esteem										0
Outcomes										
Depression		0	0	+		+	0	+		+
Anxiety				+		+				
Stress				+[b]			0			
Caregiver strain				+						
Caregiver burden	0		0				0	+		
Physical burden								0		
Self-perceived health										
Quality of life		0	0		0					
Mediators										
Intention to get support				+[b]						
Social support										0
Coping				0			0[b]			
Others										
Self-efficacy				+[b]			0[b]			
Knowledge					+[b]		+[b]			
Skill application					+[b]					
Perceived bonding								+		

[a] +: Statistically significant effect ($P<.05$) of the intervention on the measured outcome (either improving positive factors or decreasing adverse factors); 0: not statistically significant effect.

[b] Validation process of the measure was not reported.

Outcomes

A list of all the outcomes measured and statistically significant effects found at time of completion for each intervention is classified according to the Stress Process Model in Table 6. Outcomes were assessed with self-reported measures in all the included studies, either online [27-29,31,34,35], during phone calls [32,33,36,37], and/or face-to-face interviews [30,32,33,38]. All studies assessed outcomes at time of completion. Some also had an assessment half-way through the intervention [28,29,31,35,36,38] and two studies had a follow-up period of one [37] and six months [30].

Concerning outcomes at time of completion, none of the Web-based education interventions reported statistically significant differences on any outcomes when compared to usual care [32,36,38]. Efficacious interventions were thus found within the self-help and human-supported Web-based therapeutic interventions categories [27-31,33-35,37]. The most frequently assessed outcome was depression; it was shown as being significantly decreased by four interventions [27-29,31,37]. The intervention by Beauchamp et al [27] generated the largest number of positive effects on caregivers; in addition to decreased depression and anxiety, the intervention had positive effects on intrapsychic strains (eg, increase in caregiver gain), on mediators of stress (eg, intention to get support), and on self-efficacy. Namkoong et al [35] found that users of their multicomponent intervention, which integrated peer support, experienced a sense of bonding with the other participants, which in turn had a positive influence on mediators of stress such as coping abilities. In terms of longitudinal outcomes, neither of the two studies with a long-term follow-up showed statistically significant outcomes at these time points [30,37].

Discussion

The goal of this study was to systematically review the best available evidence regarding the efficacy of Internet-based interventions for caregivers of older adults. Specifically, we sought to narratively synthesize the components integrated in such interventions following the classification of Barack and Klein [19] and the taxonomy of Abraham and Michie [22] to eventually link intervention components and BCTs with outcomes on caregivers' stress process and well-being. Twelve studies with a high level of evidence covering 10 Web-based Internet interventions were found and analyzed in depth. A synthesis of the results in a comprehensive table is available online (Multimedia Appendix 2).

Categories and Components of Internet-Based Interventions

Results from the review concerned Web-based education intervention, self-help Web-based therapeutic interventions, and human-supported Web-based therapeutic interventions. Online counseling, Internet-operated therapeutic software (including emerging technologies such as robotics, therapeutic gaming, and three-dimensional environments), and other online activities were not found in studies of a high level of evidence, which may reflect the novelty of research in these categories. Studies from these categories of Internet-based intervention are currently either at a pilot stage or have a lower level of evidence [39,54], which were not considered in this review.

Concerning the components, a combination of interactive online activities and provision of human support seemed to generally lead to better outcomes in caregivers. Exercises, homework, and questionnaires were the most used components from the interactive online activities' category and appeared to be part of the success of the efficacious interventions. This can be explained by the fact that they linked to the use of BCTs and to the provision of human support. On one hand, exercises and homework were used to reinforce and build on the knowledge and skills caregivers learned while reading or viewing the content of the interventions, which can be viewed as the usage of "prompt practice" and "model behavior" techniques. In this way, Internet-based interventions represent a valuable advantage over telephone-based interventions or printed educational material [17] because they can enable participation to such interactive online exercises. On the other hand, results of online questionnaires were often sent to health professionals, which contributed to creating a continuous link between participants and clinicians. The ability to easily communicate with care providers and to monitor one's own health or the health of the care recipient can be of high importance for remote caregivers, especially if they do not live with the care recipient [68]. Effective monitoring, with online questionnaires and planned professional support, also addresses a need for longitudinal assessment of caregiver outcomes, as recommended by guidelines for interventions for caregivers [7]. Furthermore, interactive online activities were used to tailor the content accessed by caregiver, thus rendering the intervention more meaningful and personal to every participant's own needs and situation. Tailoring is an effective way to transmit content and to engage participants in an intervention [69,70], which can be easily done by Internet-based interventions with internal algorithms. Overall, interactive online activities may be used as an axle component that link different parts of the intervention to make it more appealing and engaging for the caregiver.

Human support, either provided by a health professional or peers, asynchronously or synchronously, was a component widely used in the interventions included and might account for the reported efficacy of human-supported Web-based therapeutic interventions. Having rapid and remote access to a health professional for advice and tailored support has been reported in previous studies of Internet-based interventions for care recipients as the primary factor predicting adherence [69,71] and positive effects on behavior change [20]. Qualitative results from previous research on Internet usage for professional support among caregivers suggest remote professional support was appreciated for the rapidity of the answers [47] and can alleviate barriers caregivers sometimes experience during face-to-face appointments (eg, feeling like a burden for the professional, hesitant to ask questions or express feelings) [72]. Knowing that they have access at any time and place to professional support may also make caregivers feel less worried [47,72]. Support from a group of peers was also found as a key factor of the efficacy of Internet-based interventions for older adults [71] and is reported as one of the three main reasons of using Internet among caregivers of people with cancer [73]. Furthermore, Kinnane and Milne [73] reported that caregivers viewed online support groups as means to communicate information about caring, planning for future steps to come, exchanging support in difficult moments, and venting feelings. Thus, peer support appears to be an effective component of BCT, possibly in relation to the fact that it can help support the development of effective coping strategies, which are important mediators of stress [9]. It is difficult to say if the way human support is provided (asynchronous vs synchronous) has any impact on the efficacy of human support itself because most of the included interventions used asynchronous human support. Asynchronous communication can cost less to developers and be used more easily by caregivers than complex synchronous communication modes, such as videoconference. However, in a study comparing a chat group to a video support group for caregivers of people with Alzheimer disease, caregivers reported feeling more at ease on the video group and experienced a more natural communication with the other caregivers, despite technical difficulties [44]. This sense of closeness and fluent conversation should not be understated and could have potential impact on caregiver outcomes, especially in interventions focusing on providing peer support. Overall, available evidence and results from this review suggest that human support should be considered as an efficacious component when designing Internet-based interventions.

Behavior Change Techniques

Despite recommendations [20], integration of BCTs was not consistent across interventions. Indeed, little more than half of the 26 possible BCTs [22] were incorporated in the Internet-based interventions and concerned mostly the provision of "social support," "instructions," and "problem-solving techniques." Yet, in a meta-analysis of Internet-based interventions promoting behavior change, Webb et al [20] found that interventions incorporating stress management had the greatest effect on behavior. Only four interventions used this technique in our sample [28-31,34,35] by offering caregivers concrete strategies to reduce stress, relaxation activities, or a detailed action plan to manage daily life and stressors. Moreover, results from the most recent national survey on caregiving in the United States show that stress management is the second information caregivers want to obtain, after information on how to maintain the care recipient at home [3]. Lastly, none of the three Web-based education interventions, which included only generic education content, had significant effects on any caregiver outcomes [32,36,38], which has also been reported by a review of Internet-based interventions for caregivers of people with dementia [11]. This can be explained by the lack of behavior change content and the use of very few to no BCTs to support the caregiver in achieving better health and well-being. Likely, traditional education interventions, delivered through printed material or face-to-face, have been shown to have little to no effect on caregivers' outcomes [74]. In general, Internet-based interventions should incorporate more BCTs with possibly more focus on stress management techniques and contributing factors.

Outcomes

The efficacy of Internet-based interventions for caregivers of older adults has been primarily demonstrated for psychological outcomes, such as a reduction in depression, anxiety, and burden, in this review and previous ones [11-14,16,17]. These results, although informative, permit only partial understanding of the underlying causes on which the interventions act. Indeed, in the Stress Process Model, depression, anxiety, and burden are only the outcomes of a long chain of primary and secondary stressors. Very few of the included interventions sought to measure more proximal indicators of well-being in the stress process, such as the quality of the relationship with the care recipient, disruptiveness, self-efficacy, and sense of mastery. Yet, these indicators are of clinical interest [7] and can prevent overexhaustion and serious mental difficulties if they are targeted earlier in the service delivery continuum. Furthermore, Internet-based interventions should be more concerned with enhancing mediators of stress, such as coping strategies and social support, as these are among the principal predictors of burden and decreased health for caregivers of older adults [75]. Finally, studies that measured quality of life did not show significant differences between groups at three months [34,36,38], possibly because these outcomes may not improve during a relatively short intervention.

Research Gaps and Recommendations

Methodological and reporting differences in the studies limit the conclusions that can be currently drawn concerning Internet-based interventions for caregivers of older adults. Firstly, there was considerable heterogeneity in factors that can impact on the efficacy of the intervention, namely the dosage, the adherence, and the comfort of the users with the Internet. Only half of the studies reported usage metrics or adherence statistics, making it difficult to establish the frequency and the length of usage needed for an Internet-based intervention to reach full efficacy. Without knowing if participants adhered to instructions given, it is also difficult to draw conclusions about the feasibility of the intervention or whether it adds to the burden of care. To rectify this situation, intended use should be described and adherence should be carefully monitored throughout each trial and reported as a process outcome [69], with reasons for nonadherence detailed. This could inform future studies to determine the best delivery protocols for Internet-based interventions.

Secondly, chosen control conditions for all studies, except two [29,37], did not permit to exactly isolate which components were associated with efficacy. Rather, we can only hypothesize on which combination of components and factors can lead to better outcomes in caregivers. Although comparing Internet-based interventions to usual care or educational e-bulletin serves to demonstrate their combined efficacy, it does not permit isolation of the effects of each component specifically. The work of Chih et al [29] is a good example of how to obtain that information; the participants in the control group had access to the same Internet-based intervention (CHESS), but without the component of interest (eg, clinical reports). Isolation of the effects of each component could help future research provide better cost-benefit analyses because some components require more resources than others to develop. This could also help developers and decision makers in health care prioritize what components to incorporate in Internet-based interventions to maximize the efficacy of current services.

Thirdly, there were several methodological weaknesses within the studies that reduced the validity of their findings, namely lack of blinding, high rates of attrition, uncontrolled risks for co-intervention, and unclear reporting. Blinding is a difficult criterion to fulfill within psychosocial intervention trials, but not impossible as demonstrated by Blom et al [28]. Future Internet-based trials should do likewise to reduce bias in reported outcomes. Unsurprisingly, trials that targeted caregivers of people with terminal or degenerative conditions (eg, cancer, dementia) registered higher levels of attrition, mainly due to the death of the care recipient or the overwhelming burden of the caregiver. Some studies did not clearly report critical information, which made it difficult to determine the validity of methods employed. Following guidelines for reporting Internet-based trials, such as the CONSORT-EHEALTH guidelines [25], could improve reporting and increase the confidence in the findings.

Finally, with the current sample it is not possible to determine if positive effects of Internet-based interventions are maintained over time. Indeed, measures were mainly taken on completion of programs and only two studies had a moderate to long-term follow-up period (one and six months) [30,37]. Future trials could document long-term outcomes to compare the efficacy of Internet-based interventions to traditional face-to-face interventions over time.

Limitations and Strengths

In terms of methods, there are several limits to this systematic review. First, a meta-analysis of the data was not performed given the heterogeneity of the outcomes, which restricts the findings to the state of hypotheses. Second, initially included studies were not counterverified by a second author. To ensure that we would capture the best evidence available despite this weakness, the research strategy was expanded to all possible wording of keywords of interest. Uncertainty concerning the inclusion of a study was always resolved with a second or third opinion. Therefore, we are confident that this systematic review covers the best evidence currently available in the field of Internet-based interventions for caregivers of older adults. Third, coding of the included interventions was performed by only one author and relied on the information reported by the authors in the studies, which might not adequately represent all the components of the delivered intervention. A hallmark of this review was the use of appropriate categorizations to describe and analyze the Internet-based interventions and the use of BCTs. This proved relevant in comparing different interventions with varying levels of interactivity and guidance, which helped to draw useful conclusions.

Conclusions

The findings from this systematic review suggest that Internet-based interventions with tailored behavior change content that are interactive, provide human support either by professionals or peers, and incorporate BCTs, such as provision of specific instructions regarding the behavior, problem solving, and stress management, can have positive effects on the psychological well-being of caregivers of older adults. Further randomized controlled trials that demonstrate the effect of each component individually with appropriate control conditions, analyze their outcomes considering adherence to protocol, and structure their report according to reporting guidelines in eHealth are needed to strengthen the validity of these results.

Acknowledgments

Funding for this project was provided from grant AGEWELL NCE CRP 2015-WP2.2 with salary support from Fonds de recherche en santé du Québec (Auger, Ahmed) and Canadian Institutes of Health Research (Mortenson, Miller).

The team wishes to thank Pamela Harrison, biomedical librarian, who helped in developing the research strategy of this review.

Conflicts of Interest

None declared.

Multimedia Appendix 1

MEDLINE MeSH terms and keywords.

[PDF File (Adobe PDF File), 36KB - jmir_v19i9e313_app1.pdf]

Multimedia Appendix 2

Synthesis of components, behaviour change techniques and outcomes for each category of Internet-based interventions.

[PDF File (Adobe PDF File), 51KB - jmir_v19i9e313_app2.pdf]

References

1. Walker A, Pratt C, Eddy L. Informal caregiving to aging family members: a critical review. Fam Relat 1995;44(4):402-411. [doi: 10.2307/584996]
2. Turcotte M, Sawaya C. Statistique Canada. Soins aux aînés: différences selon le type de logement URL: http://www.statcan.gc.ca/pub/75-006-x/2015001/article/14142-fra.pdf [accessed 2017-04-20] [WebCite Cache ID 6psGobeGI]
3. National Alliance for Caregiving, AARP Public Policy Institute. Caregivers of Older Adults: A Focused Look at Those Caring for Someone Age 50+. 2015 Jun. URL: http://www.caregiving.org/wp-content/uploads/2015/05/2015_CaregivingintheUS_Care-Recipients-Over-50_WEB.pdf [accessed 2017-04-20] [WebCite Cache ID 6psGz39ET]
4. Karlamangla A, Tinetti M, Guralnik J, Studenski S, Wetle T, Reuben D. Comorbidity in older adults: nosology of impairment, diseases, and conditions. J Gerontol A Biol Sci Med Sci 2007;62(3):296-300. [doi: 10.1093/gerona/62.3.296]
5. Zarit S, Todd P, Zarit J. Subjective burden of husbands and wives as caregivers: a longitudinal study. Gerontologist 1986 Jun 01;26(3):260-266. [doi: 10.1093/geront/26.3.260]

6. Turcotte M. Statistics Canada. 2013 Sep. Family caregiving: what are the consequences? URL: http://www.statcan.gc.ca/pub/75-006-x/2013001/article/11858-eng.pdf [accessed 2017-04-20] [WebCite Cache ID 6psH8vzlO]
7. Adelman R, Tmanova L, Delgado D, Dion S, Lachs M. Caregiver burden: a clinical review. JAMA 2014 Mar 12;311(10):1052-1060. [doi: 10.1001/jama.2014.304] [Medline: 24618967]
8. Sinha M. Statistics Canada. 2013 Sep. Portrait of caregivers URL: http://www.statcan.gc.ca/pub/89-652-x/89-652-x2013001-eng.pdf [accessed 2017-04-20] [WebCite Cache ID 6psGUKi55]
9. Pearlin L, Mullan J, Semple S, Skaff M. Caregiving and the stress process: an overview of concepts and their measures. Gerontologist 1990 Oct 01;30(5):583-594. [doi: 10.1093/geront/30.5.583]
10. Navaie M. Accessibility of caregiver education and support programs: reaching hard-to-reach caregivers. In: Toseland RW, Haigler DH, Monahan DJ, editors. Education and Support Programs for Caregivers. New York: Springer New York; 2011:13-28.
11. 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]
12. Chi N, Demiris G. A systematic review of telehealth tools and interventions to support family caregivers. J Telemed Telecare 2015;21(1):37-44. [doi: 10.1177/1357633×14562734]
13. Hu C, Kung S, Rummans TA, Clark MM, Lapid MI. Reducing caregiver stress with internet-based interventions: a systematic review of open-label and randomized controlled trials. J Am Med Inform Assoc 2015 Apr;22(e1):e194-e209 [FREE Full text] [doi: 10.1136/amiajnl-2014-002817] [Medline: 25125686]
14. Kaltenbaugh DJ, Klem ML, Hu L, Turi E, Haines AJ, Hagerty LJ. Using Web-based interventions to support caregivers of patients with cancer: a systematic review. Oncol Nurs Forum 2015 Mar;42(2):156-164. [doi: 10.1188/15.ONF.156-164] [Medline: 25806882]
15. Magnusson L, Hanson E, Nolan M. The impact of information and communication technology on family carers of older people and professionals in Sweden. Ageing Soc 2005 Aug 23;25(05):693. [doi: 10.1017/S0144686X05003673]
16. Perkins E, LaMartin K. The Internet as social support for older carers of adults with intellectual disabilities. J Policy Pract Intel 2012;9(1):53-62. [doi: 10.1111/j.1741-1130.2012.00330.x]
17. Wu Y, Faucounau V, de Rotrou J, Riguet M, Rigaud A. [Information and communication technology interventions supporting carers of people with Alzheimer's disease: a literature review]. Psychol Neuropsychiatr Vieil 2009;7(3):185-192. [doi: 10.1684/pnv.2009.0175]
18. Lewis B, Williams D, Neighbors C, Jakicic J, Marcus B. Cost analysis of Internet vs print interventions for physical activity promotion. Psychol Sport Exerc 2010;11(3):246-249. [doi: 10.1016/j.psychsport.2009]
19. Barak A, Klein B, Proudfoot JG. Defining internet-supported therapeutic interventions. Ann Behav Med 2009 Aug;38(1):4-17. [doi: 10.1007/s12160-009-9130-7] [Medline: 19787305]
20. Webb TL, Joseph J, Yardley L, Michie S. Using the internet to promote health behavior change: a systematic review and meta-analysis of the impact of theoretical basis, use of behavior change techniques, and mode of delivery on efficacy. J Med Internet Res 2010 Feb;12(1):e4 [FREE Full text] [doi: 10.2196/jmir.1376] [Medline: 20164043]
21. Murray E. Web-based interventions for behavior change and self-management: potential, pitfalls, and progress. Med 2 0 2012 Aug;1(2):e3 [FREE Full text] [doi: 10.2196/med20.1741] [Medline: 25075231]
22. Abraham C, Michie S. A taxonomy of behavior change techniques used in interventions. Health Psychol 2008 May;27(3):379-387. [doi: 10.1037/0278-6133.27.3.379] [Medline: 18624603]
23. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 2009 Jul 21;6(7):e1000097 [FREE Full text] [doi: 10.1371/journal.pmed.1000097] [Medline: 19621072]
24. Schünemann H, Brozek J, Guyatt G, Oxman A. Handbook for Grading the Quality of Evidence and the Strength of Recommendations Using the GRADE Approach. 2013. URL: http://gdt.guidelinedevelopment.org/app/handbook/handbook.html [WebCite Cache ID 6psHGGOvH]
25. Eysenbach G, CONSORT E-Health Group. CONSORT-EHEALTH: improving and standardizing evaluation reports of Web-based and mobile health interventions. J Med Internet Res 2011 Dec;13(4):e126 [FREE Full text] [doi: 10.2196/jmir.1923] [Medline: 22209829]
26. Higgins J, Green S. Cochrane Handbook for Systematic Reviews of Interventions Version 5. 2011. URL: http://handbook.cochrane.org [accessed 2017-04-20] [WebCite Cache ID 6psHOBmms]
27. Beauchamp N, Irvine AB, Seeley J, Johnson B. Worksite-based Internet multimedia program for family caregivers of persons with dementia. Gerontologist 2005 Dec 01;45(6):793-801. [doi: 10.1093/geront/45.6.793]
28. Blom M, Zarit S, Groot ZR, Cuijpers P, Pot A. Effectiveness of an internet intervention for family caregivers of people with dementia: results of a randomized controlled trial. PLoS ONE 2015;10(2):a. [doi: 10.1371/journal.pone.0116622]
29. Chih M, Dubenske L, Hawkins R, Brown R, Dinauer S, Cleary J. Communicating advanced cancer patients' symptoms via the Internet: a pooled analysis of two randomized trials examining caregiver preparedness, physical burden, and negative mood. Palliative Med 2013;27(6):533-543. [doi: 10.1177/0269216312457213]

30. Cristancho-Lacroix V, Wrobel J, Cantegreil-Kallen I, Dub T, Rouquette A, Rigaud AS. A web-based psychoeducational program for informal caregivers of patients with Alzheimer's disease: a pilot randomized controlled trial. J Med Internet Res 2015 May 12;17(5):e117 [FREE Full text] [doi: 10.2196/jmir.3717] [Medline: 25967983]
31. DuBenske L, Gustafson D, Namkoong K, Hawkins R, Atwood A, Brown R, et al. CHESS improves cancer caregivers' burden and mood: results of an eHealth RCT. Health Psychol 2014 Oct;33(10):1261-1272 [FREE Full text] [doi: 10.1037/a0034216] [Medline: 24245838]
32. Eames S, Hoffmann T, Worrall L, Read S, Wong A. Randomised controlled trial of an education and support package for stroke patients and their carers. BMJ Open 2013 May 08;3(5):e002538 [FREE Full text] [doi: 10.1136/bmjopen-2012-002538] [Medline: 23657469]
33. Kim J, Lee S, Kim J. Effects of a web-based stroke education program on recurrence prevention behaviors among stroke patients: a pilot study. Health Educ Res 2013;28(3):488-501. [doi: 10.1093/her/cyt044]
34. McLaughlin KA, Glang A, Beaver SV, Gau JM, Keen S. Web-based training in family advocacy. J Head Trauma Rehabil 2013;28(5):341-348 [FREE Full text] [doi: 10.1097/HTR.0b013e31824e1d43] [Medline: 22688211]
35. Namkoong K, DuBenske LL, Shaw BR, Gustafson DH, Hawkins RP, Shah DV, et al. Creating a bond between caregivers online: effect on caregivers' coping strategies. J Health Commun 2012;17(2):125-140 [FREE Full text] [doi: 10.1080/10810730.2011.585687] [Medline: 22004055]
36. Pierce LL, Steiner VL, Khuder SA, Govoni AL, Horn LJ. The effect of a Web-based stroke intervention on carers' well-being and survivors' use of healthcare services. Disabil Rehabil 2009;31(20):1676-1684. [doi: 10.1080/09638280902751972] [Medline: 19479528]
37. Smith G, Egbert N, Dellman-Jenkins M, Nanna K, Palmieri P. Reducing depression in stroke survivors and their informal caregivers: a randomized clinical trial of a web-based intervention. Rehabil Psychol 2012;57(3):196-206. [doi: 10.1037/a0029587]
38. Torkamani M, McDonald L, Saez AI, Kanios C, Katsanou M, Madeley L, ALADDIN Collaborative Group. A randomized controlled pilot study to evaluate a technology platform for the assisted living of people with dementia and their carers. J Alzheimers Dis 2014;41(2):515-523. [doi: 10.3233/JAD-132156] [Medline: 24643137]
39. Chiang L, Chen W, Dai Y, Ho Y. The effectiveness of telehealth care on caregiver burden, mastery of stress, and family function among family caregivers of heart failure patients: a quasi-experimental study. Int J Nurs Stud 2012;49(10):1230-1242. [doi: 10.1016/j.ijnurstu.2012.04.013]
40. Czaja SJ, Loewenstein D, Schulz R, Nair SN, Perdomo D. A videophone psychosocial intervention for dementia caregivers. Am J Geriatr Psychiatry 2013 Nov;21(11):1071-1081. [doi: 10.1016/j.jagp.2013.02.019] [Medline: 23831174]
41. 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]
42. Kajiyama B, Thompson LW, Eto-Iwase T, Yamashita M, Di Mario J, Marian TY, et al. Exploring the effectiveness of an internet-based program for reducing caregiver distress using the iCare Stress Management e-Training Program. Aging Ment Health 2013 Jul;17(5):544-554 [FREE Full text] [doi: 10.1080/13607863.2013.775641] [Medline: 23461355]
43. Marziali E, Donahue P. Caring for others: Internet video-conferencing group intervention for family caregivers of older adults with neurodegenerative disease. Gerontologist 2006;46(3):398-403. [doi: 10.1093/geront/46.3.398]
44. Marziali E, Garcia LJ. Dementia caregivers' responses to 2 Internet-based intervention programs. Am J Alzheimers Dis Other Demen 2011 Feb;26(1):36-43. [doi: 10.1177/1533317510387586] [Medline: 21282276]
45. Pagán-Ortiz ME, Cortés DE, Rudloff N, Weitzman P, Levkoff S. Use of an online community to provide support to caregivers of people with dementia. J Gerontol Soc Work 2014 Jul;57(6-7):694-709 [FREE Full text] [doi: 10.1080/01634372.2014.901998] [Medline: 24689359]
46. van der Roest HG, Meiland FJ, Jonker C, Dröes R. User evaluation of the DEMentia-specific Digital Interactive Social Chart (DEM-DISC). A pilot study among informal carers on its impact, user friendliness and, usefulness. Aging Ment Health 2010 May;14(4):461-470. [doi: 10.1080/13607860903311741] [Medline: 20455122]
47. 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]
48. Dew M, Goycoolea J, Harris R, Lee A, Zomak R, Dunbar-Jacob J. An internet-based intervention to improve psychosocial outcomes in heart transplant recipients and family caregivers: development and evaluation. J Heart Lung Transplant 2004;23(6):745-758. [doi: 10.1016/j.healun.2003.07.002]
49. Ducharme F, Dube 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 Nurs Inform J 2011;6(2):e1.
50. Glueckauf R, Ketterson T, Loomis J, Dages P. Online support and education for dementia caregivers: overview, utilization, and initial program evaluation. Telemed J E Health 2004;10(2):223-232. [doi: 10.1089/1530562041641354]
51. Klemm PR, Hayes ER, Diefenbeck CA, Milcarek B. Online support for employed informal caregivers: psychosocial outcomes. Comput Inform Nurs 2014 Jan;32(1):10-20. [doi: 10.1097/CIN.0000000000000009] [Medline: 24284908]

52. Kwok T, Au A, Wong B, Ip I, Mak V, Ho F. Effectiveness of online cognitive behavioral therapy on family caregivers of people with dementia. Clin Interv Aging 2014 Apr;9:631. [doi: 10.2147/CIA.S56337] [Medline: 24748781]
53. Lewis ML, Hobday JV, Hepburn KW. Internet-based program for dementia caregivers. Am J Alzheimers Dis Other Demen 2010 Dec 04;25(8):674-679. [doi: 10.1177/1533317510385812]
54. Lexis M, Everink I, van Der Heide L, Spreeuwenberg M, Willems C, de Witte L. Activity monitoring technology to support homecare delivery to frail and psychogeriatric elderly persons living at home alone. Technol Disabil 2013;25:189-197. [doi: 10.3233/TAD-130377]
55. Lorig K, Thompson-Gallagher D, Traylor L, Ritter PL, Laurent DD, Plant K, et al. Building better caregivers. J Appl Gerontol 2010 Dec 15;31(3):423-437. [doi: 10.1177/0733464810389806]
56. Magnusson L, Hanson E, Brito L, Berthold H, Chambers M, Daly T. Supporting family carers through the use of information and communication technology—the EU project ACTION. Int J Nurs Stud 2002 May;39(4):369-381. [doi: 10.1016/S0020-7489(01)00034-7]
57. Mahoney D, Mutschler P, Tarlow B, Liss E. Real world implementation lessons and outcomes from the Worker Interactive Networking (WIN) project: workplace-based online caregiver support and remote monitoring of elders at home. Telemed J E Health 2008;14(3):224-234. [doi: 10.1089/tmj.2007.0046]
58. McKechnie V, Barker C, Stott J. The effectiveness of an Internet support forum for carers of people with dementia: a pre-post cohort study. J Med Internet Res 2014 Feb;16(2):e68 [FREE Full text] [doi: 10.2196/jmir.3166] [Medline: 24583789]
59. Nahm E, Resnick B, Orwig D, Magaziner J, Bellantoni M, Sterling R, et al. A theory-based online hip fracture resource center for caregivers: effects on dyads. Nurs Res 2012;61(6):413-422 [FREE Full text] [doi: 10.1097/NNR.0b013e318270712a] [Medline: 23052420]
60. Northouse L, Schafenacker A, Barr KL, Katapodi M, Yoon H, Brittain K, et al. A tailored web-based psychoeducational intervention for cancer patients and their family caregivers. Cancer Nurs 2014;37(5):1. [doi: 10.1097/NCC.0000000000000159]
61. Sander AM, Clark AN, Atchison TB, Rueda M. A web-based videoconferencing approach to training caregivers in rural areas to compensate for problems related to traumatic brain injury. J Head Trauma Rehabil 2009;24(4):248-261. [doi: 10.1097/HTR.0b013e3181ad593a]
62. Scott K, Beatty L. Feasibility study of a self-guided cognitive behaviour therapy Internet intervention for cancer carers. Aust J Prim Health 2013;19(4):270. [doi: 10.1071/PY13025]
63. Stjernswärd S, Östman M. Illuminating user experience of a website for the relatives of persons with depression. Int J Soc Psychiatry 2010 Mar 16;57(4):375-386. [doi: 10.1177/0020764009358388]
64. Torp S, Hanson E, Hauge S, Ulstein I, Magnusson L. A pilot study of how information and communication technology may contribute to health promotion among elderly spousal carers in Norway. Health Soc Care Community 2008 Jan;16(1):75-85. [doi: 10.1111/j.1365-2524.2007.00725.x] [Medline: 18181817]
65. Lundberg S. The results from a two-year case study of an information and communication technology support system for family caregivers. Disabil Rehabil Assist Technol 2014 Jul;9(4):353-358. [doi: 10.3109/17483107.2013.814170] [Medline: 23855387]
66. Sabat S. Flourishing of the self while caregiving for a person with dementia: a case study of education, counseling, and psychosocial support via email. Dementia 2011;10(1):91-97. [doi: 10.1177/1471301210392986]
67. Williams K, Arthur A, Niedens M, Moushey L, Hutfles L. In-home monitoring support for dementia caregivers: a feasibility study. Clin Nurs Res 2013;22(2):139-150. [doi: 10.1177/1054773812460545]
68. Zulman DM, Piette JD, Jenchura EC, Asch SM, Rosland A. Facilitating out-of-home caregiving through health information technology: survey of informal caregivers' current practices, interests, and perceived barriers. J Med Internet Res 2013 Jul 10;15(7):e123 [FREE Full text] [doi: 10.2196/jmir.2472] [Medline: 23841987]
69. 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]
70. 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]
71. 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;16(Suppl 5):290. [doi: 10.1186/s12913-016-1522-3]
72. Schook RM, Linssen C, Schramel FM, Festen J, Lammers E, Smit EF, et al. Why do patients and caregivers seek answers from the Internet and online lung specialists? A qualitative study. J Med Internet Res 2014 Feb;16(2):e37 [FREE Full text] [doi: 10.2196/jmir.2842] [Medline: 24496139]
73. Kinnane NA, Milne DJ. The role of the Internet in supporting and informing carers of people with cancer: a literature review. Support Care Cancer 2010 Sep;18(9):1123-1136. [doi: 10.1007/s00520-010-0863-4] [Medline: 20336326]
74. Selwood A, Johnston K, Katona C, Lyketsos C, Livingston G. Systematic review of the effect of psychological interventions on family caregivers of people with dementia. J Affect Disord 2007;101(1-3):75-89. [doi: 10.1016/j.jad.2006]

75. Bień-Barkowska K, Doroszkiewicz H, Bień B. Silent strain of caregiving: exploring the best predictors of distress in family carers of geriatric patients. Clin Interv Aging 2017 Feb;12:263-274 [FREE Full text] [doi: 10.2147/CIA.S125664] [Medline: 28203067]

Abbreviations

BCTs: behavior change techniques
CONSORT: Consolidated Standards of Reporting Trials
GRADE: Grading of Recommendations Assessment, Development and Evaluation
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses

Edited by G Eysenbach; submitted 20.04.17; peer-reviewed by F Routhier, K Best, L Boots; comments to author 28.06.17; revised version received 11.07.17; accepted 11.07.17; published 19.09.17.

Please cite as:
Guay C, Auger C, Demers L, Mortenson WB, Miller WC, Gélinas-Bronsard D, Ahmed S
Components and Outcomes of Internet-Based Interventions for Caregivers of Older Adults: Systematic Review
J Med Internet Res 2017;19(9):e313
URL: http://www.jmir.org/2017/9/e313/
doi:10.2196/jmir.7896
PMID:28928109

Viewpoint

A Dermatologist's Ammunition in the War Against Smoking: A Photoaging App

Titus Josef Brinker[1,2,3], MD; Alexander Enk[1], MD; Martina Gatzka[4], MD; Yasuhiro Nakamura[5], MD; Wiebke Sondermann[2,3,6], MD; Albert Joachim Omlor[7], MD; Maximilian Philip Petri[2,3,6], MD; Ante Karoglan[8], MD; Werner Seeger[9], MD, PhD; Joachim Klode[2,3,6], MD, PhD; Christof von Kalle[10,11,12], MD; Dirk Schadendorf[2,3,6], MD, PhD

[1]Department of Dermatology and National Center for Tumor Diseases (NCT), University Hospital Heidelberg, University of Heidelberg, Heidelberg, Germany

[2]Department of Dermatology, Venereology and Allergology, University-Hospital Essen, University of Duisburg-Essen, Essen, Germany

[3]German Cancer Consortium, Heidelberg, Germany

[4]Department of Dermatology and Allergic Diseases, University of Ulm, Ulm, Germany

[5]Department of Skin Oncology/Dermatology, Saitama Medical University International Medical Center, Saitama, Japan

[6]West German Cancer Center, University of Duisburg-Essen, Essen, Germany

[7]Department of Experimental Pneumology and Allergology, Saarland University Faculty of Medicine, Homburg, Germany

[8]Department of Dermatology, University Hospital Magdeburg, University of Magdeburg, Magdeburg, Germany

[9]Department of Internal Medicine, Universities of Giessen and Marburg Lung Center; Member of the German Center for Lung Research, Justus-Liebig-University of Giessen, Gießen, Germany

[10]National Center for Tumor Diseases (NCT), University of Heidelberg, Heidelberg, Germany

[11]Heidelberg Center for Personalized Oncology (DKFZ-HIPO), Heidelberg, Germany

[12]German Cancer Research Center (DKFZ) and German Cancer Consortium (DKTK), Division of Translational Oncology, University of Heidelberg, Heidelberg, Germany

Corresponding Author:
Titus Josef Brinker, MD
Department of Dermatology and National Center for Tumor Diseases (NCT)
University of Heidelberg
Im Neuenheimer Feld 440/460
Heidelberg,
Germany
Phone: 49 151 7508 4347
Email: titus.brinker@gmail.com

Abstract

This viewpoint reviews the perspectives for dermatology as a specialty to go beyond the substantial impact of smoking on skin disease and leverage the impact of skin changes on a person's self-concept and behavior in the design of effective interventions for smoking prevention and cessation.

(J Med Internet Res 2017;19(9):e326) doi:10.2196/jmir.8743

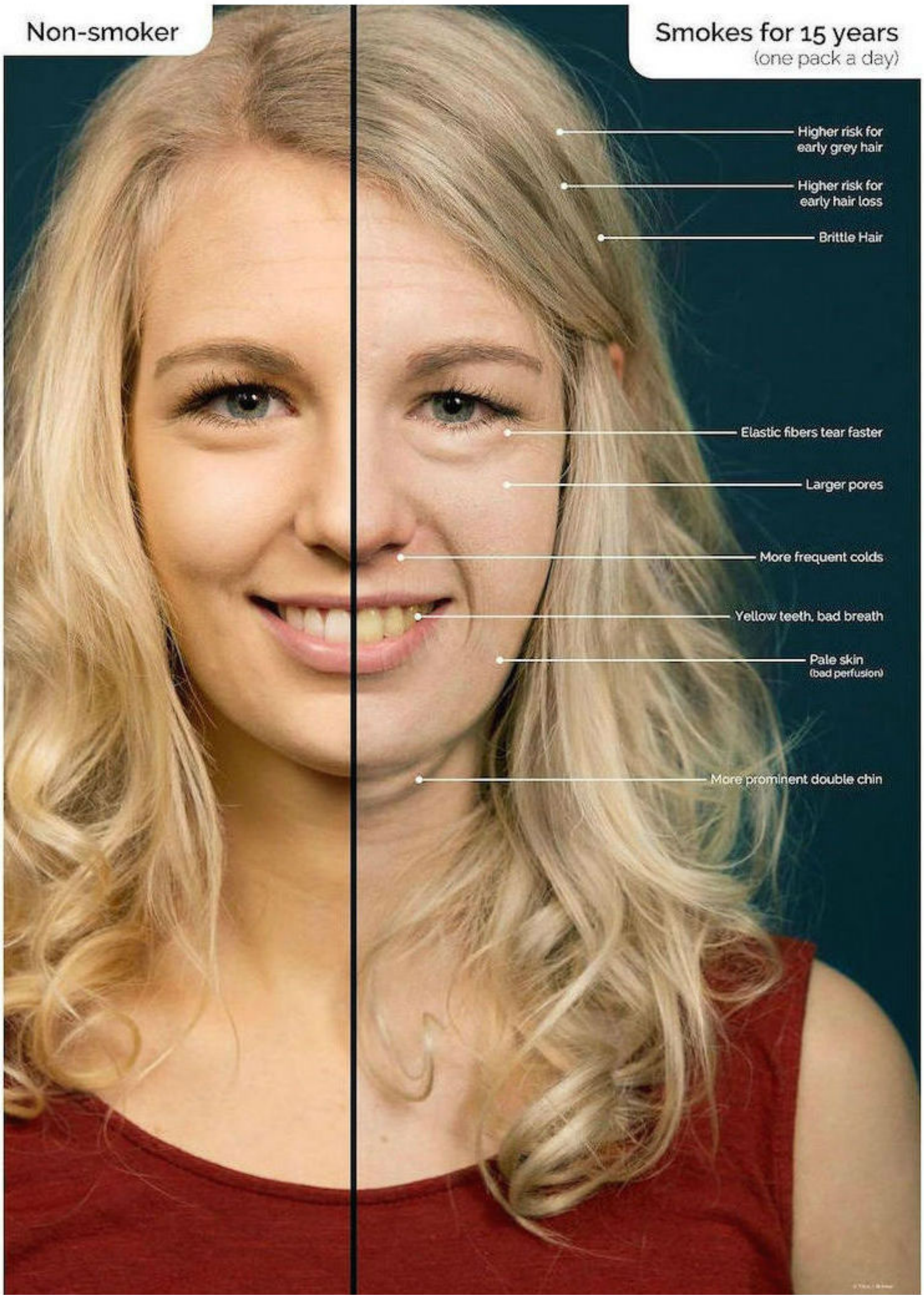
KEYWORDS

dermatology; smoking; apps; photoaging; face; skin; tobacco; tobacco cessation; tobacco prevention

Most smokers start smoking during their early adolescence, often with the idea that smoking is glamorous; the problems related to impaired wound healing, erectile dysfunction, and oral cancers are too far in the future to fathom. In contrast, for the majority of teenagers, attractiveness is the most important predictor of their own self-esteem [1].

Interventions focusing on the negative dermatologic changes due to smoking have been effective in altering behavior, both in adolescence [2-4] and young adulthood [5,6]. Skin damage due to smoking that is culturally associated with a decrease in attractiveness (ie, wrinkles, early hair loss, declined capillary perfusion, pale or grayish skin [7-9]) predominantly affects the self-concept of young people with low education [1], who are at significantly greater risk for tobacco addiction [10-12] and benefit the most from abstinence [13]. After reviewing the evidence regarding facial changes due to smoking on PubMed, we designed Figure 1 in order to extrapolate the typical appearance of a smoker's face as frequently seen and noted by dermatologists.

Figure 1. Normal aging versus effects of smoking a pack a day for 15 years.



First steps have been taken to disseminate this dermatologic knowledge on irreversible aesthetic damage to the target groups and measure its effectiveness in randomized trials (ie, via the free photoaging app Smokerface, in which a selfie is altered to predict future appearance) in Germany [3,4,14,15] and Brazil [16] with a total of more than 150,000 downloads. In addition, photoaging desktop-based interventions in France [6], Switzerland [2], and Australia [5] showed promising results that justify definitive randomized trials. The relevance of skin-based appearance for individual behavior was also confirmed in the setting of skin cancer prevention [4,17-21].

Dermatology as an interdisciplinary specialty needs to go beyond the substantial impact of smoking on skin disease [22,23] and leverage the impact of skin changes on a person's self-concept [1] and behavior [5] in the design of effective interventions for the largest cause of preventable death and disease in the western world [24]. Future dermatologic research should focus on developing, evaluating, and optimizing new ways to implement the specialty's superior ammunition in the war against smoking.

References

1. Baudson T, Weber K, Freund PA. More than only skin deep: appearance self-concept predicts most of secondary school students' self-esteem. Front Psychol 2016;7:1568 [FREE Full text] [doi: 10.3389/fpsyg.2016.01568] [Medline: 27803681]
2. Weiss C, Hanebuth D, Coda P, Dratva J, Heintz M, Stutz EZ. Aging images as a motivational trigger for smoking cessation in young women. Int J Environ Res Public Health 2010 Dec;7(9):3499-3512 [FREE Full text] [doi: 10.3390/ijerph7093499] [Medline: 20948939]
3. Brinker T, Owczarek A, Seeger W, Groneberg D, Brieske C, Jansen P, et al. A medical student-delivered smoking prevention program, education against tobacco, for secondary schools in Germany: randomized controlled trial. J Med Internet Res 2017 Jun 06;19(6):e199 [FREE Full text] [doi: 10.2196/jmir.7906] [Medline: 28588007]
4. Brinker TJ, Seeger W, Buslaff F. Photoaging mobile apps in school-based tobacco prevention: the mirroring approach. J Med Internet Res 2016 Jun 28;18(6):e183 [FREE Full text] [doi: 10.2196/jmir.6016] [Medline: 27352819]
5. Burford O, Jiwa M, Carter O, Parsons R, Hendrie D. Internet-based photoaging within Australian pharmacies to promote smoking cessation: randomized controlled trial. J Med Internet Res 2013 Mar 26;15(3):e64 [FREE Full text] [doi: 10.2196/jmir.2337] [Medline: 23531984]
6. Burford O, Kindarji S, Parsons R, Falcoff H. Using visual demonstrations in young adults to promote smoking cessation: preliminary findings from a French pilot study. Res Social Adm Pharm 2017 May 04:1. [doi: 10.1016/j.sapharm.2017.04.050] [Medline: 28495124]
7. Okada H, Alleyne B, Varghai K, Kinder K, Guyuron B. Facial changes caused by smoking: a comparison between smoking and nonsmoking identical twins. Plast Reconstr Surg 2013 Nov;132(5):1085-1092. [doi: 10.1097/PRS.0b013e3182a4c20a] [Medline: 23924651]
8. Dupati A, Helfrich YR. Effect of cigarette smoking on skin aging. Expert Rev Dermatol 2009;4(4):371-378.
9. Yin L, Morita A, Tsuji T. Skin aging induced by ultraviolet exposure and tobacco smoking: evidence from epidemiological and molecular studies. Photodermatol Photoimmunol Photomed 2001 Aug;17(4):178-183. [Medline: 11499540]
10. Cutler DM, Lleras-Muney A. Understanding differences in health behaviors by education. J Health Econ 2010 Jan;29(1):1-28 [FREE Full text] [doi: 10.1016/j.jhealeco.2009.10.003] [Medline: 19963292]
11. Hoebel J, Kuntz B, Kroll L, Finger J, Zeiher J, Lange C, et al. Trends in absolute and relative educational inequalities in adult smoking since the early 2000s: the case of Germany. Nicotine Tob Res 2017 Apr 18:1. [doi: 10.1093/ntr/ntx087] [Medline: 28431153]
12. Kuntz B, Lampert T. Smoking and passive smoke exposure among adolescents in Germany: prevalence, trends over time, and differences between social groups. Dtsch Arztebl Int 2016;113(3):23-30.
13. Doll R, Peto R, Boreham J, Sutherland I. Mortality in relation to smoking: 50 years' observations on male British doctors. BMJ 2004;328(7455):50.
14. Brinker TJ, Holzapfel J, Baudson TG, Sies K, Jakob L, Baumert HM, et al. Photoaging smartphone app promoting poster campaign to reduce smoking prevalence in secondary schools: the Smokerface Randomized Trial: design and baseline characteristics. BMJ Open 2016 Nov 07;6(11):e014288 [FREE Full text] [doi: 10.1136/bmjopen-2016-014288] [Medline: 27821601]
15. Brinker T, Seeger W. Photoaging mobile apps: a novel opportunity for smoking cessation? J Med Internet Res 2015;17(7):e186.
16. Xavier L, Bernardes-Souza B, Lisboa O, Seeger W, Groneberg D, Tran T, et al. A medical student-delivered smoking prevention program, Education Against Tobacco, for secondary schools in Brazil: study protocol for a randomized trial. JMIR Research Protocols 2017 2017;6(1):e16.
17. Brinker TJ, Schadendorf D, Klode J, Cosgarea I, Rösch A, Jansen P, et al. Photoaging mobile apps as a novel opportunity for melanoma prevention: pilot study. JMIR Mhealth Uhealth 2017 Jul 26;5(7):e101 [FREE Full text] [doi: 10.2196/mhealth.8231] [Medline: 28747297]
18. Olson AL, Gaffney CA, Starr P, Dietrich AJ. The impact of an appearance-based educational intervention on adolescent intention to use sunscreen. Health Educ Res 2008 Oct;23(5):763-769 [FREE Full text] [doi: 10.1093/her/cym005] [Medline: 18039727]
19. Owen A, Grogan S, Clark-Carter D. Effects of an appearance-focussed versus a health-focussed intervention on men's attitudes towards UV exposure. Int J Men Health 2016;15(1):34.
20. Tuong W, Armstrong AW. Effect of appearance-based education compared with health-based education on sunscreen use and knowledge: a randomized controlled trial. J Am Acad Dermatol 2014;70(4):665-669. [Medline: 24508292]
21. Brinker T, Brieske CM, Schaefer CM, Buslaff F, Gatzka M, Petri MP, et al. Photoaging Mobile Apps in School-Based Melanoma Prevention: Pilot Study. J Med Internet Res 2017 Sep 08;19(9):e319 [FREE Full text] [doi: 10.2196/jmir.8661] [Medline: 28887295]
22. Ortiz A, Grando SA. Smoking and the skin. Int J Dermatol 2012 Mar;51(3):250-262. [doi: 10.1111/j.1365-4632.2011.05205.x] [Medline: 22348557]

23. Dusingize J, Olsen C, Pandeya N, Subramaniam P, Thompson B, Neale R, et al. Cigarette smoking and the risks of basal cell carcinoma and squamous cell carcinoma. J Invest Dermatol 2017 Aug;137(8):1700-1708. [doi: 10.1016/j.jid.2017.03.027] [Medline: 28414022]
24. GBD 2015 Tobacco Collaborators. Smoking prevalence and attributable disease burden in 195 countries and territories, 1990-2015: a systematic analysis from the Global Burden of Disease Study 2015. Lancet 2017 May 13;389(10082):1885-1906 [FREE Full text] [doi: 10.1016/S0140-6736(17)30819-X] [Medline: 28390697]

Edited by G Eysenbach; submitted 14.08.17; peer-reviewed by O Burford, T Langley; comments to author 30.08.17; revised version received 30.08.17; accepted 30.08.17; published 21.09.17.

Please cite as:
Brinker TJ, Enk A, Gatzka M, Nakamura Y, Sondermann W, Omlor AJ, Petri MP, Karoglan A, Seeger W, Klode J, von Kalle C, Schadendorf D
A Dermatologist's Ammunition in the War Against Smoking: A Photoaging App
J Med Internet Res 2017;19(9):e326
URL: http://www.jmir.org/2017/9/e326/
doi:10.2196/jmir.8743
PMID:28935619

Original Paper

Effectiveness of a Web-Based Tailored Intervention With Virtual Assistants Promoting the Acceptability of HPV Vaccination Among Mothers of Invited Girls: Randomized Controlled Trial

Mirjam Pot[1,2], PhD; Theo GWM Paulussen[1], PhD; Robert AC Ruiter[2], Prof Dr; Iris Eekhout[1,3], PhD; Hester E de Melker[4], PhD; Maxine EA Spoelstra[5], MSc; Hilde M van Keulen[1], PhD

[1]Netherlands Organization for Applied Scientific Research (TNO), Child Health, Leiden, Netherlands
[2]Department of Work and Social Psychology, Maastricht University, Maastricht, Netherlands
[3]VU University Medical Center, Epidemiology & Biostatistics, Amsterdam, Netherlands
[4]National Institute for Public Health and the Environment (RIVM), Centre for Infectious Disease Control, Bilthoven, Netherlands
[5]Department of Psychology, Leiden University, Leiden, Netherlands

Corresponding Author:
Mirjam Pot, PhD
Netherlands Organization for Applied Scientific Research (TNO)
Child Health
Schipholweg 77
Leiden, 2316 ZL
Netherlands
Phone: 31 643234293
Email: mirjampot90@gmail.com

Related Article:

This is a corrected version. See correction statement: https://www.jmir.org/2020/7/e22565/

Abstract

Background: In 2010, the human papillomavirus (HPV) vaccination was introduced in the Dutch National Immunization Program for 12-year-old girls, aiming to reduce the incidence of cervical cancer in women. HPV vaccination uptake turned out to be lower than expected: 61% versus 70%, respectively. Mothers were shown to play the most important role in the immunization decision about this vaccination. They had also expressed their need for interactive personal information about the HPV vaccination over and above the existing universal general information. To improve the effectiveness of the existing education about the HPV vaccination, we systematically developed a Web-based tailored intervention with virtual assistants providing mothers of girls to be invited with tailored feedback on their decision making about the HPV vaccination.

Objective: The aim of this study was to evaluate the effectiveness of the Web-based tailored intervention for promoting HPV vaccination acceptance by means of a randomized controlled trial (RCT).

Methods: Mothers were recruited via the Dutch vaccination register (Praeventis) (n=36,000) and three Web-based panels (n=2483). Those who gave informed consent (N=8062) were randomly assigned to the control (n=4067) or intervention condition (n=3995). HPV vaccination uptake, as registered by Praeventis once the HPV vaccination round was completed, was used as the primary outcome. Secondary outcomes were differential scores across conditions between baseline (before the provided access to the new tailored intervention) and follow-up (just before the first vaccination) regarding the mothers' degree of informed decision making (IDM), decisional conflict, and critical determinants of HPV vaccination uptake among which are intention, attitude, risk perception, and outcome beliefs.

Results: Intention-to-treat analysis (N=8062) showed a significant positive effect of the intervention on IDM, decisional conflict, and nearly all determinants of HPV vaccination uptake (*P*<.001). No effect was found on uptake (*P*=.60). This may be attributed to the overall high uptake rates in both conditions. Mothers evaluated the intervention as highly positive, including the website as well as the virtual assistants that were used to deliver the tailored feedback.

Conclusions: This computer-tailored intervention has the potential to improve HPV vaccination acceptability and IDM and to decrease decisional conflict among mothers of invited girls. Implications for future research are discussed.

Trial Registration: Trialregister.nl NTR4935; http://www.trialregister.nl/trialreg/admin/rctview.asp?TC=4935 (Archived by WebCite at http://www.webcitation.org/6srT7l9EM)

(J Med Internet Res 2017;19(9):e312) doi:10.2196/jmir.7449

KEYWORDS

vaccination; health promotion; web-based intervention computer-tailoring; randomized controlled trial

Introduction

Worldwide, cervical cancer is the third most common cancer in women [1]. Persistent infection by the human papillomavirus (HPV) is the causative agent of cervical cancer [2]. In the Netherlands, yearly 750 new cases of cervical cancer are detected, of which 242 are with fatal consequences [3], despite the presence of a national cervical cancer screening program for women aged 30 to 60 years [4]. In 2008, the Health Council advised the Ministry to include the HPV vaccination for girls aged 12 years in the National Immunization Program (NIP) [5]. Initial implementation started with a catch-up campaign in 2009 for girls aged 13 to 16 years. From 2010, new cohorts of 12-year-old girls have been invited by the NIP to receive the HPV vaccination on a yearly basis. The municipal health services organize local sessions for group-based HPV vaccination, usually at large venues. This restricts the opportunity for interaction between the parent and girl with the professional. The vaccination itself is given by young health professionals (ie, medical doctors, nurses, or doctor's assistants). The vaccination is voluntary and is offered free of charge. Complete vaccination includes 2 injections with a 6-month interval.

So far, HPV vaccination uptake in the Netherlands has remained lower than expected: 61% uptake in 2016 while 70% was targeted [6]. Research has indicated that mothers play the most important role in decision making about the girls' immunization [7]. Currently, the regular invitation for the HPV vaccination comprises an introduction folder and a link to a website providing universal information about HPV and HPV vaccination. However, research has already indicated that mothers feel more in need for interactive personal information about the HPV vaccination, over and above this universal information [8]. To improve the existing educational strategy targeting HPV vaccination uptake, we developed a computer-tailored intervention with virtual assistants using the 6-step Intervention Mapping protocol for developing theory- and evidence-based health promotion interventions [9]. The intervention was aimed at Dutch mothers of girls to be invited for the HPV vaccination in 2015 (ie, girls born in 2002).

To date, only few tailored interventions to encourage HPV vaccination have been tested [10-13]. Three of these showed positive results. Hopfer [11] found that HPV vaccination uptake doubled among participants who were exposed to a culturally tailored video, compared with controls. Gerend and colleagues [12] found that, compared with general information, information tailored to the individual participant's perceived barriers increased HPV vaccination intention. Grandahl and colleagues [13] found that an intervention delivered individually, that is face-to-face, by school nurses positively affected beliefs toward HPV prevention as well as vaccination uptake. However, to our knowledge, effective interventions promoting HPV vaccination that can reach large groups at relatively low costs (eg, Web-based tailored interventions) [14] do not exist yet.

Social cognitive determinants of the mothers' decision making about their daughters' HPV vaccination that appeared both relevant and changeable were selected as targets for developing the intervention [9]. These included HPV vaccination-related intention, attitude, outcome beliefs, risk perception, anticipated regret, subjective norms, habit, relative effectiveness of the HPV vaccination, and self-efficacy [7,8,15]. These determinants appeared to account for large proportions of variance in the mothers' decision-making outcome (80-82%) [7]. Also, large proportions of the mothers do not actively acquire and process information about the pros and cons of the HPV vaccination and feel ambivalent about their decision [7,8]. This indicates that these decisions are based on rather unstable grounds, which makes them vulnerable for arguments challenging their initial attitudes and/or intention. Because informed decision making (IDM) is expected to make mothers less vulnerable for counterarguments, this was also chosen as a relevant intervention target. According to Marteau and colleagues [16], an informed decision is based in sufficient and relevant knowledge and in the congruence between the person's values (ie, their attitude toward the HPV vaccination) and the behavioral outcome (ie, whether mothers had their daughters vaccinated against HPV). Consequently, knowledge was also targeted by the intervention. In addition, decisional conflict was selected as a target as this appears strongly related to IDM, with the possibility that decisional conflict may arise when feeling uninformed [17].

The aim of this study was to assess the effectiveness of the Web-based tailored intervention with virtual assistants on HPV vaccination uptake among the participants' daughters (primary outcome). Secondary outcomes were the mothers' degree of IDM, decisional conflict, and the social cognitive determinants of decision making about the daughters' HPV vaccination uptake (eg, attitude, intention, and beliefs). When compared with the control condition, significantly positive effects were expected in the intervention condition with respect to HPV vaccination uptake, social cognitive determinants of the mothers' decision making about the vaccination, levels of informed decision making, and levels of decisional conflict.

Methods

Participants

Mothers were randomly recruited from Praeventis, the Dutch National Immunization Register, and three Internet panels. The latter was to assure a high response rate [7]. This gave us the opportunity to assess differential effects of the intervention under (1) more controlled conditions (ie, panel sample) and (2) more naturalistic conditions as will be the case when the intervention has become part of the national implementation strategy (ie, Praeventis sample). This provided a basis for inferences concerning the intervention's efficacy and effectiveness [18]. This is also why we did not reward the mothers from the Praeventis sample. The panel members received a small financial reimbursement for each survey that they completed (2-3 euros per survey). The amount of money received was panel specific. Extra financial reimbursement was provided to those in the intervention condition who completed all surveys (1-3 euros extra). In total, the participants in the control group could receive between 4 and 6 euros (completing baseline and follow-up), whereas participants in the experimental group could receive between 5 and 9 euros (completing baseline, intervention, and follow-up). Panel members were prestratified by region to ensure geographic diversity.

Power Calculation

Power analyses indicated that 1200 mothers per sample (ie, Praeventis and panels) were needed at baseline (600 per arm) to detect a 10% difference in HPV vaccination uptake between the intervention and control group, and small effects on the continuous secondary outcomes (Cohen d=0.10-0.30), with a power of 0.80, a two-sided alpha of .05, and an expected dropout of 30% at the last survey. Given previous experiences [7,8], a response rate of 3% was expected in the Praeventis sample. A total of 36,000 mothers were initially invited to participate via Praeventis and 1200 mothers via the Internet panels.

Design

This study, approved by the Ethical Committee of the VU Medical Center in Amsterdam, was conducted between January 2015 and March 2015. Effectiveness was evaluated by a 2-arm randomized controlled trial (RCT). Intervention effects on HPV vaccination uptake were assessed objectively using Praeventis. Because invited girls were given the opportunity to catch up on their missed HPV vaccinations, complete data on uptake were only available 18 months after baseline (ie, July 2016). The effects on secondary outcomes were examined by two Web-based surveys; at baseline, just before they had access to the experimental education, and at follow-up, just before they received the first HPV vaccination (time intervals around 2 months). Participants in the intervention condition were invited to visit the Web-based intervention between baseline and follow-up. Participants in both arms had access to the universal information about the HPV vaccination as part of the regular invitation for the HPV vaccination.

Procedure

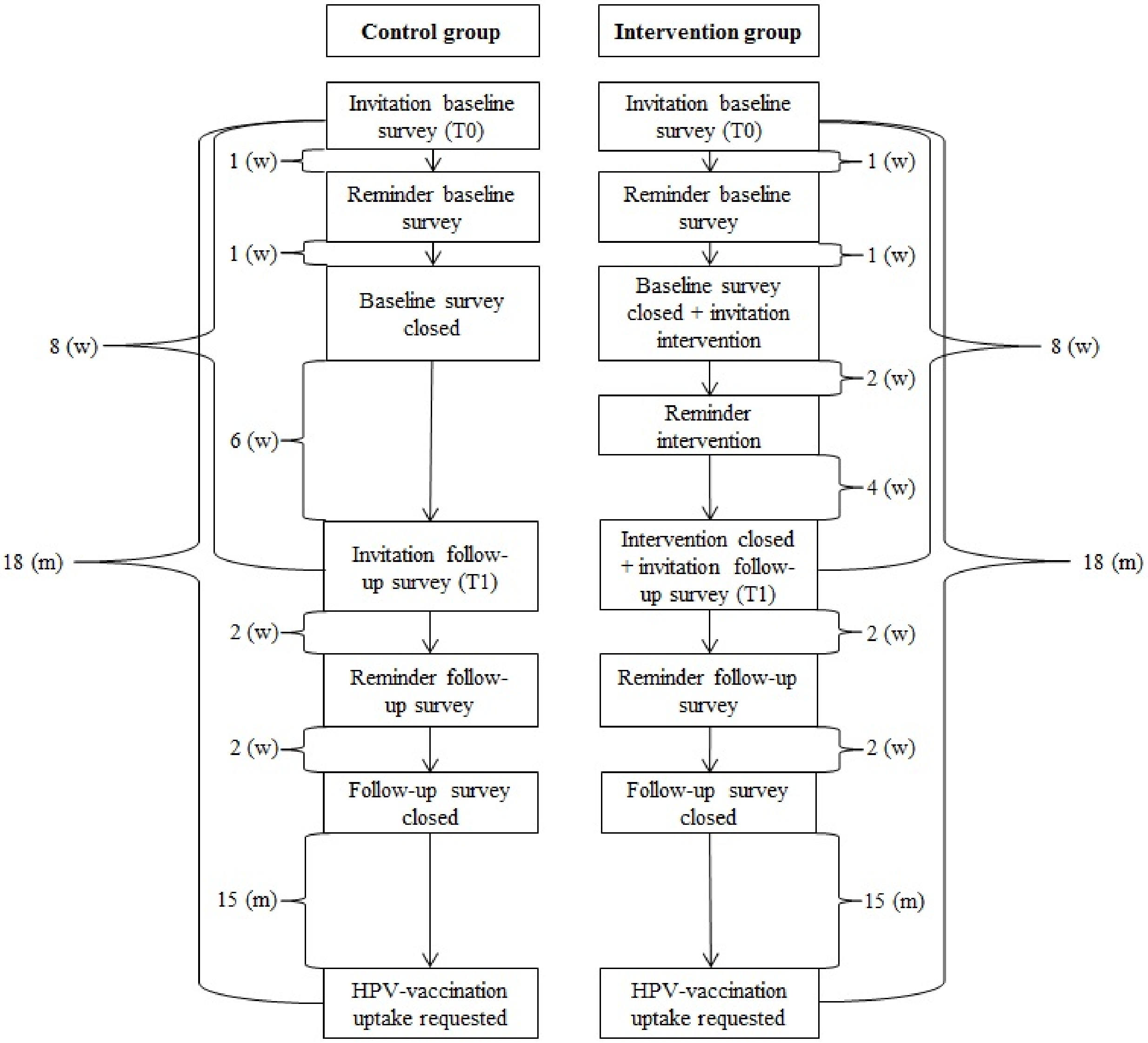
An invitation letter to participate in the study was sent in January 2015 by postal mail to the Praeventis sample and by email to the panel sample. This letter included information about the study, a link to a secured website, and a unique code for entrance to the baseline survey. The same code was used for entering the follow-up survey and for gaining access to the tailored intervention (only participants in the intervention condition). Securing the website and providing unique codes was done to reduce the risk of spillover effects. The mothers in the intervention condition were explicitly requested not to share the link with others. On the website, the participants were assured of their privacy, confidentiality, and security in handling their responses and were informed that they could withdraw from participation at any time. Participants were then asked to provide informed consent and to give us permission to derive their daughters' HPV vaccination status from Praeventis. After having provided informed consent, participants were randomly assigned to either the control or intervention condition. A reminder was sent 1 week after the first invitation. One week after the reminder, participants in the intervention condition received an email inviting them to visit the Web-based tailored feedback. Two weeks after this invitation, a reminder was sent. The website could be visited until the invitation to complete the follow-up questionnaire, 8 weeks after the initial invitation for the baseline questionnaire. All participants were given 2 weeks to complete the follow-up survey; a reminder was sent after the first week. These timelines (see Figure 1) were chosen as these fit with the standard procedures for the HPV vaccination in the Netherlands.

Intervention

The intervention consisted of a website providing mothers with tailored feedback from 2 virtual assistants, one being visualized in Figure 2. In Multimedia Appendix 1, a selection of screenshots of the website is presented. Computer-tailoring was the basic method for change and fitted the outcome of a previously conducted needs assessment indicating that the mothers preferred personalized feedback [8]. Tailoring is a health communication strategy in which messages are individualized to the person's preferences and needs [19]. Meta-analyses have shown computer-tailored interventions to be more effective than universal interventions in achieving behavioral change outcomes (eg, [20,21]). Examples of change techniques that were used in addition to tailoring were consciousness raising (targeting risk perception) [22], belief selection (targeting beliefs) [23], active learning (targeting knowledge retention) [24], and motivational interviewing (targeting decisional conflict and attitude change) [25]. Before completing the concept intervention, we experimentally pretested three different intervention components [26,27] and conducted three focus groups for pretesting the prototype's feasibility.

Also, innovative was the use of 2 virtual assistants for delivering the tailored feedback; a mother- and doctor-like assistant. They provided opportunities for two-way interactions and for creating a highly personal experience. The added value of using a virtual assistant over a text and/or picture-based website is that it improves information recall [28], transfer of learning [29], amount of learning [30], self-efficacy expectations, literacy, and behavioral change [31-33].

Figure 1. Timeline of data collection for the control and intervention group.



The website consisted of four menu options: (1) two-sided information about the HPV vaccination, (2) a decisional balance, (3) practical background information, and (4) frequently asked questions (see Multimedia Appendix 1). In the first menu, mothers were able to collect tailored information about the HPV vaccination such as information about the risk of contracting an HPV infection, which may cause cervical cancer, as well as the risks and effectiveness of the HPV vaccine. In the second menu, a decisional balance gave mothers the opportunity to weigh their perceived pros and cons to balance the mothers' position toward vaccinating versus not vaccinating the daughter (see Figure 2). In the third menu, mothers received practical information such as how and where to get the HPV vaccination and how to talk to their daughter and/or partner about the HPV vaccination.

Outcome Measures

Primary Outcome: HPV Vaccination Uptake

HPV vaccination uptake was derived from Praeventis, which was registered as having received no, 1, or 2 injections. We dichotomized HPV vaccination uptake into having received no HPV injection (0=not vaccinated) versus having received 1 or 2 HPV injections (1=vaccinated), as data showed that the determinants of HPV vaccination contrasted in these groups contrasted the most.

Secondary Outcomes: IDM and Decisional Conflict

According to Marteau and colleagues [16], a choice is considered to be informed when people have sufficient and relevant knowledge (knowledge) and when the person's values (attitude) and behavior match. As such, IDM is usually expressed dichotomously (eg, see [34]). However, as we think the selection of the cut-off points is somewhat arbitrary, yet critical for the outcome [35], we also constructed a continuous measure for IDM. Post hoc analyses showed the correlation to be high between the two (Spearman ρ [rho]=.78). Both scores were derived from the Multi-dimensional Measure of Informed Choice [16,34,36].

Figure 2. Screenshot of the decisional balance and the mother-like virtual assistant on the website.

Stelling	Uw antwoorden			Hoe belangrijk is dit voor u?
Mijn dochter is te jong om ingeënt te worden tegen HPV.	[✓ Eens]	Neutraal	Oneens	★★☆
Als mijn dochter zich laat inenten tegen HPV, dan zal zij in de toekomst onveilig vrijen.	Eens	Neutraal	[✓ Oneens]	★☆☆
Als de overheid meisjes laat inenten tegen HPV dan ga ik er vanuit dat de inenting veilig is voor mijn dochter.	[✓ Eens]	Neutraal	Oneens	☆☆☆
Mijn dochter is nog niet seksueel actief, dus hoeft ze van mij de HPV-inenting niet te krijgen.	Eens	Neutraal	[✓ Oneens]	★☆☆
Mijn dochter is al een tijd lang seksueel actief is, dus heeft het geen zin meer om haar te laten inenten tegen HPV.	Eens	[✓ Neutraal]	Oneens	☆☆☆
De HPV-inenting verkleint de kans dat mijn dochter later baarmoederhalskanker krijgt.	[✓ Eens]	Neutraal	Oneens	★★★
De meeste moeders in mijn omgeving laten hun dochter ook inenten tegen HPV.	Eens	[✓ Neutraal]	Oneens	☆☆☆
Mijn dochter wil zich niet laten inenten.	Eens	Neutraal	[✓ Oneens]	☆☆☆
De persoon die meebeslist wil niet dat	Eens	Neutraal	[✓ Oneens]	★☆☆

By constructing the dichotomous IDM, the mother was classified as an informed decision maker, if (1) she had sufficient knowledge and (2) her attitude was consistent with the behavior (HPV vaccination uptake). Knowledge was considered to be sufficient when it was higher than the baseline mean score. A decision was considered to be consistent when the mother's attitude was positive (higher than 4 on a 7-point scale) and the daughter was vaccinated, or when her attitude was negative (lower than 4 on a 7-point scale) and her daughter was not vaccinated. Any other combination was categorized as an uninformed decision.

By constructing the continuous measure of IDM, we first recoded the attitude scale into a scale ranging from −3 (extremely negative) to 3 (extremely positive). Consistency (ranging from −3 to 3) resulted from multiplying this attitude score with uptake (−1=no injection vs 1=1 or 2 injections). Then, the resulting consistency score was recoded from 0 (least consistent) to 6 (most consistent). Next, knowledge (−8 to 8, see Table 1) was recoded into a scale ranging from 0 (least knowledgeable) to 8 (most knowledgeable); the original scores below 0 were all recoded to 0. Finally, the continuous scale for IDM was computed by multiplying knowledge (range 0-8) with consistency (range 0-6), resulting in a scale ranging from 0 (not/least informed) to 48 (most informed).

Decisional conflict was measured using the "Uncertainty" subscale of the Decisional Conflict Scales [37], which included three 7-point scaled items (eg, "as regards to the HPV-vaccination, I felt sure about my choice) (1=completely disagree to 7=completely agree). Internal consistency was high (Cronbach alpha=.94).

Determinants of HPV Vaccination Uptake

Composite scores were computed for determinants of HPV vaccination uptake in case the scaled items showed internal consistency (Cronbach alpha >.60 or Pearson r>.64). See Table 1 for an overview of the primary and secondary outcome measures. All scores on scaled items were averaged into a scale because they showed sufficient internal consistency (Cronbach alpha ≥.78/Pearson r ≥.64); Cronbach alpha was used for scales consisting of more than 2 items, whereas Pearson r was used for scales consisting of 2 items; items with an (R) were reverse coded.

Table 1. Overview of primary and secondary outcome measures.

Measures and items	Score/scale (minimum to maximum value)	Cronbach alpha (α) or Pearson *r* $(r)^2$	References
Primary outcome			
HPV [a] vaccination uptake	0=not vaccinated, 1=vaccinated	N/A[b]	
Uptake of the HPV vaccination is obtained through data from Praeventis.			
Secondary outcomes			
IDM [c] outcome (dichotomous)	0=not informed, 1=informed	N/A	[16], [34], [36]
An informed decision has been made when:			
The knowledge score was higher or equal to the mean of knowledge at baseline, the attitude score was higher than 4 (positive) and the HPV vaccination has been received.			
The knowledge score was higher or equal to the mean of knowledge at baseline, the attitude score was lower than 4 (negative), and the HPV vaccination has not been received.			
Any other combination was categorized as an uninformed decision.			
IDM outcome (continuous)	0=least informed decision to 48=most informed decision	N/A	[16], [34], [38]
Attitude was recoded from 0-7 to −3 (negative) to 3 (positive attitude) and HPV uptake was recoded from 0 or 1 to −1 (no injection) or 1 (1 or 2 injections).			
Level of consistency was measured by multiplying the scores for attitude by those for HPV uptake (−3=low consistency; 3=high consistency). Consistency was then recoded into 0 (low) to 6 (high). Both consistency and sufficient knowledge were considered prerequisite for an informed decision.			
Knowledge scores (−8=low; 8=high) lower or equal to 0 were considered insufficient (0=no insufficient knowledge; 8=high knowledge).			
The level of IDM outcome was determined by multiplying the scores for knowledge with those for consistency.			
Decisional conflict about the HPV vaccination	1=high to 7=low decisional conflict	.94 (α)	[17]
As regards the HPV vaccination:			
I felt sure about my choice			
The decision was relatively easy to make			
I was clear about the best choice for my daughter			
HPV vaccination intention	1=low intention to vaccinate to 7=high intention to vaccinate	.92 (*r*)	[7,8]
Are you planning on getting your daughter vaccinated against HPV?			
How big is the chance that you will get your daughter vaccinated?			
Attitude toward the HPV vaccination	1=negative to 7=positive attitude	.98 (α)	[38]
Vaccinating my daughter against HPV is...			
very undesirable to very desirable			
very bad to very good			
very negative to very positive			
very unimportant to very important			
Risk perception (having received no HPV vaccination)	1=low to 7=high risk perception	N/A	[38,39]
Imagine that your daughter was not vaccinated against HPV.			
The chance that my daughter will get cervical cancer is...			
Risk perception (having received the HPV vaccination)	1=low to 7=high risk perception	N/A	[38,39]
Imagine that your daughter was vaccinated against HPV.			
The chance that my daughter will get cervical cancer is...			

Measures and items	Score/scale (minimum to maximum value)	Cronbach alpha (α) or Pearson *r* (*r*)[2]	References
Anticipated regret about rejecting the HPV vaccination	1=low to 7=high anticipated regret	N/A	[7,8]
Imagine your daughter has not received the HPV vaccination and she gets cervical cancer in the future.			
How much would you regret your decision to let her receive no vaccination?			
Beliefs about the HPV vaccination	1=negative to 7=positive beliefs about the HPV vaccination	.85 (α)	[39,40]
If the government offers the vaccination, I assume it will be safe.			
Our government shows responsibility for the health of the Dutch population by introducing the HPV vaccination.			
The HPV vaccination was only introduced because the pharmaceutical industry will earn a lot of money from it (R).			
There is too little known about whether the HPV vaccination effectively protects against cervical cancer (R).			
There is too little known about the detrimental side effects of the HPV vaccination (R).			
My daughter is too young to receive the HPV vaccination (R).			
My daughter does not need the vaccination because she is not yet sexually active (R).			
Subjective norms toward the HPV vaccination [d]	−20=negative to 20=positive	.64 (*r*)	[38]
Normative beliefs:			
Regarding the HPV vaccination of your daughter, what is your expectation on the opinion of...			
Social referents: partner[e], daughter			
Motivation to comply:			
How motivated are you to comply with the opinion of...?			
Habit strength toward the HPV vaccination	1=weak to 7=strong habit strength	.78 (*r*)	[41]
Letting my daughter receive the HPV vaccination is something I do:			
automatically			
without thinking			
Self-efficacy expectations toward the HPV vaccination	1=low self-efficacy to 7=high self-efficacy	.82 (α)	
To what extent would you succeed in dealing with the following statements?			
Guiding my daughter in the decision regarding the HPV vaccination			
Having a good talk with my daughter about the HPV vaccination			
Having a good talk with my partner[e] about the HPV vaccination			
Motivating my daughter to have herself vaccinated			
Getting the actual HPV vaccination/2 injections with my daughter			

Measures and items	Score/scale (minimum to maximum value)	Cronbach alpha (α) or Pearson r (r)^2	References
Knowledge about the HPV vaccination [f]	−8=incorrect, 8=correct	N/A	[7,8]
Are the following statements true or false?			
HPV is sexually transmittable.			
Condoms fully protect against HPV.			
My daughter is obliged to get the HPV vaccination when she is invited.			
You will always notice when you are infected by HPV.			
Only women can get infected by HPV.			
Women who received the HPV vaccination are still advised to participate in the cervical cancer screening in the Netherlands.			
The HPV vaccination fully protects against cervical cancer.			
My daughter does not need to get the HPV vaccination if she is already sexually active.			
Relative effectiveness of the HPV vaccination [g]	−9=HPV vaccination least effective to 9=HPV vaccination most effective	N/A	[7,8]
How would you rate the effectiveness of the following methods of preventing cervical cancer:			
having safe sex			
having sex with only 1 person in a lifetime			
participating in the cervical cancer screening			
having a healthy lifestyle (eg, not smoking)			
the HPV vaccination			
Participants rated the effectiveness of each method			

[a] HPV: human papillomavirus.

[b] N/A: not applicable.

[c] IDM: informed decision making.

[d] The subjective norms score was first computed by multiplying normative beliefs and motivation to comply for each social referent, and then by summing up the multiplications of the social referents.

[e] Only applicable if the mother indicated that she had a partner.

[f] Knowledge is not a scale because the answer on 1 item does not predict the answer on other items; the items were summed up to present a sum score of knowledge.

[g] The difference between the rated effectiveness of the HPV vaccination and the most effective alternative represented the relative effectiveness score (−9=HPV vaccination least effective to 9=HPV vaccination most effective).

Sociodemographics

Sociodemographics were modeled as background variables (ie, age, educational level, country of birth, and religion). Level of education referred to the highest level the mother had completed. Educational level was classified into low (less than secondary or vocational education), intermediate (secondary through preuniversity education) or high (professional or university education) [7,8]. Country of birth was dichotomized into “Netherlands” versus “other,” as in our sample only 562 (6.97%) of 8062 (100%) mothers appeared to be born in a country other than the Netherlands. Religion was measured by asking the mothers about their religious convictions (Protestant, Roman Catholic, Muslim, Jewish, Buddhist, Hindu, other, or no religion). This was later classified as “Protestant” versus “not Protestant” as Protestants most refrain from vaccination compared with other religious or nonreligious groups in the Netherlands [7,8].

Subjective Program Evaluation and Objective Program Use

Subjective program evaluation was assessed at follow-up for mothers in the intervention condition by asking them to evaluate the website and the virtual assistants on a 10-point scale; the higher the score, the more positive the evaluation. Objective program use was evaluated by computer logs assessing the number of visits and amount of time logged in per session. If participants logged in more than once, these were summed.

Statistical Analyses

Descriptive statistics were used to describe the baseline sample. For analyzing the effects of the intervention, we used intention-to-treat (ITT) instead of complete case analysis. By using ITT, power increases while the risk of bias possibly caused by selective dropout decreases [42]. To deal with missing data, we applied multiple imputation by chained equations [42,43]. There were 15 imputed datasets generated using the predictive mean matching algorithm in Statistical Package for the Social Sciences (SPSS, IBM Corp) . The results from the imputed datasets were pooled together using Rubin's rules [37]. Convergence of the imputations was checked by inspecting the iteration plots.

Intervention effects were examined by logistic and linear regression analyses (for dichotomous and continuous variables, respectively) by using the outcome at the follow-up as the criterion and the outcome-score at baseline and condition as the independent variables [44]. In view of multiple testing, an effect was considered significant when $P<.003$ (Bonferroni corrected alpha=.05/15 factors). The odds ratio was used as an indicator for effect size (Bonferroni corrected alpha=.05/14 factors). Effect sizes for the linear regressions were calculated in R (R Development Core Team) [45] using Cohen f^2 statistic, $(R^2_{AB} – R^2_A)/(1 – R^2_{AB})$, in which B is the variable of interest (ie, condition), A is the set of all other variables (ie, the outcome at baseline), R^2_{AB} is the proportion of variance accounting for A and B together, and R^2_A is the proportion of variance accounted for by A. These were interpreted as 0.02=small, 0.15=medium, 0.35=large [46]. Complete case analyses were performed as a sensitivity check for substantial differences with the results based on ITT.

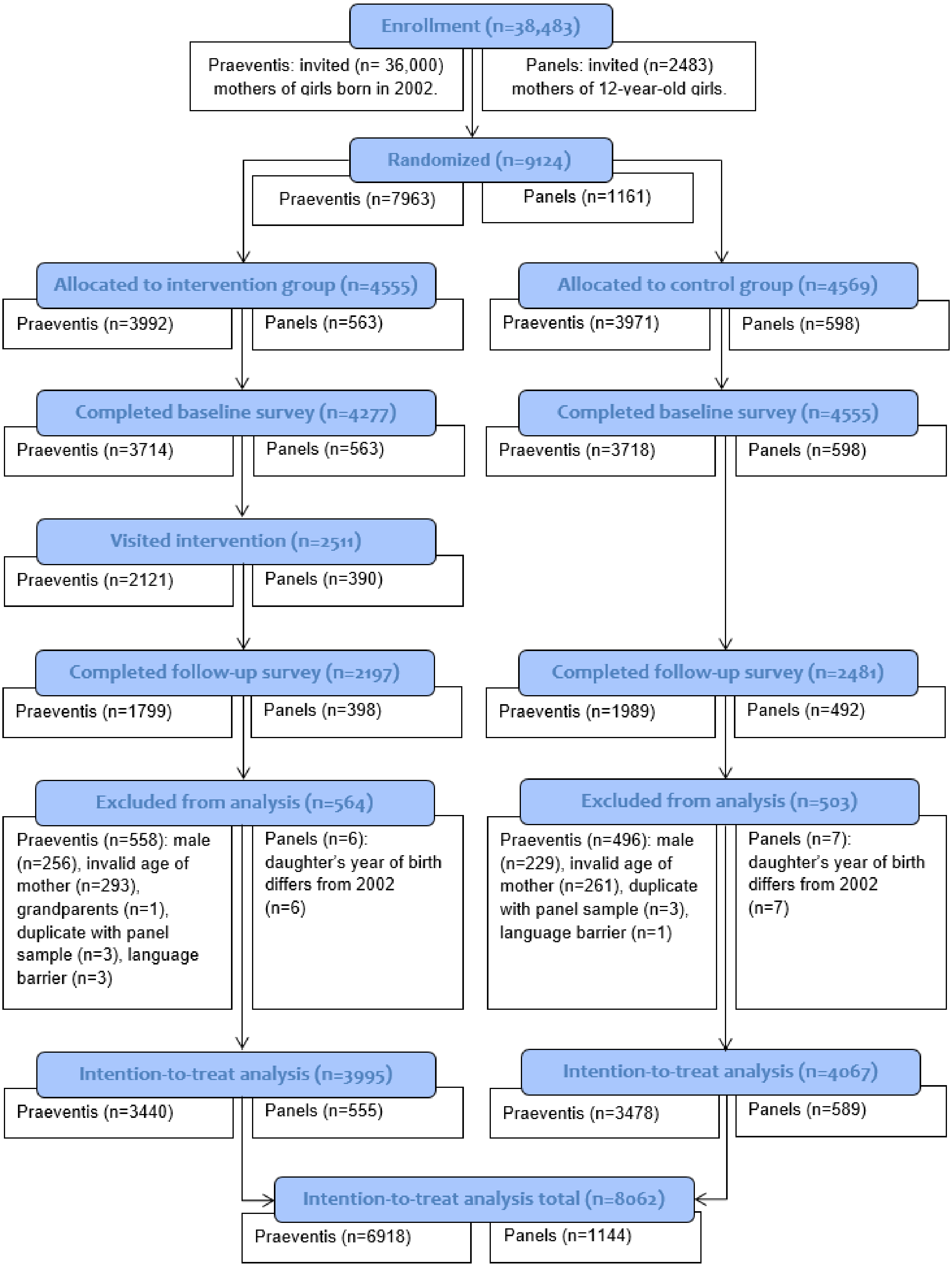
Furthermore, we performed exploratory moderation analysis (Bonferroni corrected alpha=.05/15) to examine differences in effects regarding sociodemographics (ie, age, country of birth, education level, and religion) and sample (ie, Praeventis vs panels) by including a two-way interaction term (eg, condition × sample) in each of the aforementioned regression analyses [47]. In addition, we explored whether intention (at baseline) moderated the effects found in the primary and secondary outcomes. For this, intention was divided into three subgroups: (1) mothers with a negative intention (scores below half a standard deviation (SD) below the centered mean score of intention at baseline), (2) mothers who were hesitating (scores between half an SD below and above the centered mean of intention at baseline), and (3) mothers with a positive intention (scores more than half an SD above the centered mean score). Finally, subjective evaluations and objective use of the program were assessed by using descriptive analysis. IBM statistical package SPSS version 23 was used for analyzing the data [44].

Results

Response Rates and Attrition

We invited 36,000 participants via Praeventis and 2483 via the panels. A flow diagram of the recruitment and response is shown in Figure 3. From the 9124 participants who were initially randomized at T0, 8593 (94.18%; 4277 in the intervention group and 4316 participants in the control group) completed the baseline questionnaire, whereas 4678 (51.27%; 2197 in the intervention group and 2481 participants in the control group) completed the follow-up questionnaire 8 weeks later (T1). Dropout analysis showed significantly more dropout in the Praeventis sample. There was also selective nonresponse with regard to condition (ie, more dropout in the intervention condition), sociodemographics (ie, more dropout in those not born in the Netherlands, and in those low in education), HPV vaccination uptake (more dropout in mothers having a daughter not being vaccinated), and secondary outcomes (ie, more dropout in mothers with low levels of IDM, risk perception, and self-efficacy and in mothers with high attitude scores). In total, 1067 participants were excluded (564 in the intervention group and 503 participants in the control group), as they did not meet the inclusion criteria (ie, being a mother of a daughter born in 2002 and aged 24-62 years) or were found to be duplicates across the two samples. The final sample for ITT analysis consisted of 8062 mothers: 3995 mothers in the intervention condition versus 4067 in the control condition.

Figure 3. Flow diagram of the recruitment and response of study participants. Notes: (1) Participants could be excluded based on multiple criteria (eg, a male with an invalid age). Therefore, the total number of Praeventis participants excluded differed from the sum of separate criteria for exclusion. (2) In order to assess the intervention's effectiveness (Praeventis sample) versus efficacy (panel sample), the recruitment and response is displayed per sample within each condition.



Sample Description

See Table 2 for the sample description. As there were no data available on sociodemographics of the population from which the sample was derived (ie, Dutch mothers of girls aged 12 years in the Praeventis database), we were unable to assess the representativeness of the study sample. The mean age of mothers was 43.64 years (SD 4.25). On average, mothers had a positive intention toward the HPV vaccination of their daughter at baseline (mean 5.35 [SD 1.69]). Compared with the national HPV vaccination uptake, uptake was higher in the study sample (n=59,866; 60.98% vs n=5880; 72.93%, respectively).

Table 2. Sample description (N=8062). In case of missing values, the number of missing values ($N_{missing}$) was presented. By reporting 2 decimal points for the percentages, summing the percentages for each category up may differ from 100%.

Variables	Intervention (n_{total}=3995), n (%)	Control (n_{total}=4067), n (%)	Total (N_{total}=8062), n (%)
Age in years, n (%)	43.70 (4.27)	43.58 (4.22)	43.64 (4.25)
Country of Birth, n (%)	$N_{missing}$=4 (0.10)	$N_{missing}$=4 (0.10)	$N_{missing}$=8(0.10)
The Netherlands	3715 (92.99)	3777 (92.87)	7492 (92.93)
Other	276 (6.91)	286 (7.03)	562 (6.97)
Religion, n (%)	$N_{missing}$=7 (0.18)	$N_{missing}$=6 (0.15)	$N_{missing}$=13 (0.16)
Protestant	753 (18.85)	737 (18.12)	1490 (18.48)
Not protestant	3235 (80.98)	3324 (81.73)	6559 (81.36)
Educational level, n (%)	$N_{missing}$=4 (0.10)	$N_{missing}$=3 (0.07)	$N_{missing}$=7 (0.09)
Low	588 (14.72)	540 (12.28)	1128 (13.99)
Middle	1736 (43.45)	1735 (42.66)	3471 (43.05)
High	1660 (41.55)	1786 (43.91)	3446 (42.74)
HPV[a] vaccination uptake, n (%)	2923 (73.17)	2957 (72.71)	5880 (72.93)

[a] HPV: human papillomavirus.

Intervention Effects on Primary and Secondary Outcomes

In Table 3, an overview of the intervention effects on the primary and secondary outcomes is given. A higher mean score means a higher X (eg, more positive attitude) except for decisional conflict; here, a higher mean score means less decisional conflict If an odds ratio (OR) is higher than one, this means that the higher the score on a factor, the higher the outcome of IDM or the higher the chance of the daughter being vaccinated. If an OR is less than one, this means that the higher the score on a factor, the lower outcome of IDM or the lower the chance of the daughter being vaccinated. ITT analyses showed that there was no effect of the intervention on HPV vaccination uptake (odds ratio, OR=1.03, *P*=.60). The intervention had a significant positive effect on all secondary outcomes (*P*<.001), except for risk perception when not vaccinated, anticipated regret, and self-efficacy (*P*=.01; *P*=.01; *P*=.03, respectively). Compared with the control group, at follow-up, mothers in the intervention group were more informed (dichotomous measure: OR=1.28, *P*<.001; continuous measure: beta=1.72, *P*<.003), experienced less decisional conflict (beta=.21, *P*<.003), were more intended to vaccinate their daughter (beta=.18, *P*<.004), had a more positive attitude toward vaccinating their daughter (beta=0.15, *P*<.004), had more positive beliefs (eg, beliefs about the safety and effectiveness of the HPV vaccination; beta=.12, *P*<.003), had a lower risk perception when they imagined that their daughter was vaccinated (beta=−.11, *P*<.003), perceived more positive subjective norms (beta=.82, *P*<.003), reported a higher relative effectiveness (beta=.46, *P*<.003), and had more knowledge (beta=.35, *P*<.003). Effect sizes were small (see Table 3). Results from complete case analyses were similar, except for an additional effect of the intervention on anticipated regret and self-efficacy.

Table 3. Effects of the intervention on the outcome measures according to intention-to-treat analyses (N=8062).

Outcome	Control (N=4067)		Intervention (N=3995)		Beta (standard error)	Cohen f^2 or OR
	Pretest	Posttest	Pretest	Posttest		
Primary outcome						
HPV vaccination uptake [a]						
Has received no HPV injection (reference), n (%)		1106 (27.19)		1066 (26.67)		
Has received one or two HPV injections, n (%)		2961 (72.81)		2929 (73.32)	.03 (.05)[b]	1.03
Secondary outcomes						
IDM: dichotomous						
Not informed (reference), n (%)	2689 (66.12)	1924 (47.31)	2689 (67.31)	1699 (42.53)		
Informed, n (%)	1376 (33.83)	2143 (52.69)	1306 (32.69)	2296 (57.47)	.25 (.06)[c]	1.28
IDM: continuous (0-48), mean (SD)	18.95 (11.45)	24.28 (11.82)	18.69 (11.21)	25.85 (12.30)	1.72 (.27)[c]	0.007
Decisional conflict (1-7), mean (SD)	4.33 (1.74)	5.17 (1.45)	4.33 (1.75)	5.38 (1.36)	.21 (.04)[c]	0.008
Intention (1-7), mean (SD)	5.35 (1.70)	5.42 (1.97)	5.35 (1.69)	5.59 (1.87)	.18 (.03)[c]	0.006
Attitude (1-7), mean (SD)	5.19 (1.46)	5.22 (1.57)	5.18 (1.45)	5.37 (1.51)	.15 (.03)[c]	0.006
Beliefs (1-7), mean (SD)	4.21 (.72)	4.37 (.80)	4.19 (.73)	4.47 (.81)	.12 (.02)[c]	0.010
Risk perception not vaccinated (1-7), mean (SD)	3.73 (.98)	3.70 (1.05)	3.74 (0.98)	3.77 (1.08)	.06 (.02)[d]	0.001
Risk perception vaccinated (1-7), mean (SD)	2.76 (1.06)	2.74 (1.08)	2.77 (1.07)	2.64 (1.10)	−.11 (.03)[c]	0.004
Anticipated regret (1-5), mean (SD)	3.68 (1.27)	3.50 (1.33)	3.71 (1.25)	3.59 (1.31)	.07 (.03)[d]	0.001
Subjective norm (−20 to 20), mean (SD)	5.92 (7.90)	6.46 (9.46)	5.88 (7.81)	7.25 (9.20)	.82 (.20)[c]	0.004
Habit (1-7), mean (SD)	4.26 (1.79)	4.36 (1.82)	4.28 (1.78)	4.51 (1.83)	.14 (.04)[c]	0.004
Relative effectiveness (1-10), mean (SD)	−2.01 (2.24)	−1.84 (2.36)	−1.97 (2.22)	−1.35 (2.27)	.46 (.07)[c]	0.015
Self-efficacy (1-7), mean (SD)	6.24 (.76)	6.24 (.78)	6.27 (.73)	6.29 (.75)	.04 (.02)[e]	0.001
Knowledge (−8 to 8), mean (SD)	4.42 (2.16)	5.41 (2.09)	4.40 (2.14)	5.75 (2.09)	.35 (.05)[c]	0.009

[a] Human papillomavirus (HPV) vaccination uptake was not assessed at baseline.

[b] *P*=.60.

[c] *P* ≤.001, thus significant (*P*<.003; Bonferroni: 0.05/14 factors).

[d] *P*=.01.

[e] *P*=.03.

Moderation of Intervention Effects

Regarding sociodemographics, no significant interaction effects on any of the outcome measures were found for country of birth (*P* ≥.08) or religion (*P* ≥.08). For educational level, we found an interaction effect with condition on relative effectiveness (beta=.59, *P*<.001): the intervention had more positive effects on relative effectiveness for those with high education compared with those low in education. There was no significant interaction between condition and sample on any of the outcome measures (*P* ≥.04). For the interaction effects between intention at baseline and the outcome measures, See Table 4. If an OR is higher than one, this means that the higher the score on a factor, the higher the outcome of IDM or the higher the chance of the daughter being vaccinated. If an OR is less than one, this means that the higher the score on a factor, the lower outcome of IDM or the lower the chance of the daughter being vaccinated. In the first 2 columns, the reference category is those with a negative intention. For a comparison between those in doubt (reference category) and a positive attitude, see the third column. Significant interaction effects between intention at baseline and condition were found on intention, attitude, decisional conflict, subjective norm, and relative effectiveness . For mothers who had a negative intention, the intervention had more positive effects on intention and relative effectiveness compared with mothers who were doubting (beta=.26, *P*=.002; beta=.39, *P*=.001, respectively) or had a positive intention (beta=.40, *P*<.001; beta=.53, *P*<.001, respectively). In addition, for mothers with a negative intention, the intervention had a more positive effect on attitude (beta=.21, *P*=.001) and on subjective norms (beta=1.64, *P*<.001) compared with mothers with a positive intention. For mothers who were doubting, the intervention had more positive effects on decisional conflict compared with mothers who had a negative intention (beta=.26, *P*=.001). No differences on intervention outcomes were found between mothers who were doubting and mothers who had a positive intention (*P* ≥.004).

Table 4. Moderation effects of intention subgroups on the outcome measures according to the intention-to-treat analyses (N=8062).

Outcome	Negative—in doubt		Negative—positive		In doubt—positive	
	Beta (standard error)	*P* value	Beta (standard error)	*P* value	Beta (standard error)	*P* value
Primary outcome						
HPV [a] vaccination uptake						
Has received no HPV injection (reference)						
Has received one or two HPV injections	−.21 (.13)	.22	.01 (.18)	.97	.22 (.18)	.24
Secondary outcomes						
IDM [b] : dichotomous						
Not informed (reference)						
Informed	.18 (.15)	.22	.09 (.15)	.56	−.10 (.14)	.50
IDM: continuous (0-48)	.92 (.69)	.18	1.03 (.65)	.12	.11 (.65)	.87
Decisional conflict (1-7)	.26 (.08)	.001[c]	.04 (.07)	.57	.22 (.08)	.004
Intention (1-7)	−.26 (.08)	.002[c]	−.40 (.08)	<.001[c]	−.14 (.07)	.03
Attitude (1-7)	−.17 (.06)	.009	−.21 (.06)	.001[c]	−.04 (.06)	.51
Beliefs (1-7)	−.04 (.04)	.36	−.02 (.04)	.68	.02 (.04)	.67
Risk perception not vaccinated (1-7)	.04 (.07)	.51	.09 (.07)	.18	.05 (.06)	.41
Risk perception vaccinated (1-7)	.04 (.06)	.51	−.06 (.07)	.38	−.10 (.07)	.17
Anticipated regret (1-5)	−.03 (.07)	.70	.02 (.07)	.81	.04 (.06)	.46
Subjective norm (−10 to 10)	−1.18 (.47)	.01	−1.64 (.43)	<.001[c]	−.46 (.37)	.22
Habit (1-7)	.08 (.08)	.34	.06 (.08)	.45	−.01 (.07)	.87
Relative effectiveness (1-10)	−.39 (.12)	.001[c]	−.53 (.12)	<.001[c]	−.14 (.11)	.21
Self-efficacy (1-7)	−.00 (.05)	.97	.03 (.04)	.45	.03 (.05)	.50
Knowledge (−8 to 8)	−.01 (.13)	.92	−.13 (.12)	.29	−.11 (.12)	.34

[a] HPV: human papillomavirus.

[b] IDM: informed decision making.

[c] *P*<.003, thus significant (Bonferroni: 0.05/14 factors).

Subjective Program Evaluation and Objective Program Use

Mothers in the intervention condition evaluated the website with a 7.6 (SD=1.36) and the virtual assistants with a 7.4 (SD=1.53). According to the computer logs, 2509 (62.80%) of the 3995 (100.00%) invited mothers logged on to the website. Of these, 1835 (73.14%) visited the website once, 498 (19.84%) visited twice, and 176 (7.02%) more than twice. On average, mothers spent 22 min on the website (SD=13 min).

Discussion

Principal Findings

This study investigated the effectiveness of a Web-based tailored intervention with virtual assistants promoting HPV vaccination acceptability among mothers of invited girls. As hypothesized, positive intervention effects were found with respect to the social cognitive determinants of the mothers' decision making about the vaccination (eg, HPV vaccination-related intention, attitude, and outcome beliefs), levels of IDM, and levels of decisional conflict. The positive effect of tailored education on HPV vaccination intention was also found by Gerend et al [12] among young women. However, they did not assess other determinants of HPV vaccination acceptability (next to intention) nor did they measure levels of IDM, levels of decisional conflict, or actual HPV vaccination uptake.

The findings described above suggest that this intervention has potential in promoting HPV vaccination acceptability and IDM. This is important, given the currently moderate HPV vaccination uptake and the fact that large proportions of the mothers do not actively acquire and process information about the pros and cons of this HPV vaccination and that many feel ambivalent about the decision [7,8]. Less informed decisions are decisions constituted in rather instable beliefs that are susceptible to counterarguments. Nowadays, counterarguments are all around on the Internet and Web-based social media [48]. Because the intervention initiated active processing of verifiable information about the risks and effectiveness of the HPV vaccination, it inoculates mothers with arguments that become accessible at the moment they are confronted with (new) information that might challenge their initial positive attitudes and intentions [49,50].

No effects were found on mothers' perceived risk of their daughter getting cervical cancer without the HPV vaccination, anticipated regret in case their daughter would get cervical cancer later in life, and self-efficacy. As for risk perception, baseline scores indicated that the mothers overestimated the probability of contracting cervical cancer to a great extent when taken into account the actual population incidence [51]. Because the intervention presented mothers this actual low probability of attracting cervical cancer, it seems unlikely that their perceived risk was brought to higher levels. The lack of effect on anticipated regret might be explained by the fact that we removed the intervention component specifically targeting anticipated regret. This was removed because our pilot studies and focus groups revealed that resistance was evoked by asking mothers how much regret they would have if they did not vaccinate their daughter against HPV and their daughter developed cervical cancer later in life. Also, emphasizing the impact of cervical cancer might be fear-arousing, which, in turn, may have been detrimental for exploring and processing other information provided by the program [52]. Finally, it appeared that we encountered a ceiling effect for self-efficacy as the scores at baseline among both groups were above 6 on a 7-point scale.

No effects of the tailored intervention were found on HPV vaccination uptake. This is contrary to both our expectations and to what has been found by others; Hopfer found that among female college students, HPV vaccination uptake doubled after they were exposed to a tailored video compared with controls (22% vs 12%, respectively) [11]. However, in their study, vaccination rates were quite low in the control condition (12%), leaving much room for improvement. In our study, however, the uptake rates were high in both conditions (intervention: n=3995; 73.17% and control: n=4067; 72.71%) especially when compared with the national Dutch uptake (n=59,866; 60.98) [6]. This may explain why we, as opposed to Hopfer, did not find an effect on HPV vaccination uptake. After all, we did find a larger increase in intention among mothers in the intervention compared with the control condition, and according to theory [23] and empirical findings [53], intention is an important predictor of (HPV vaccination) behavior.

Mothers evaluated the intervention as positive, specified by the high subjective evaluation of both the website (7.6 on a 10-point scale) and the virtual assistants that were used to deliver the tailored feedback (7.4 on a 10-point scale). Objective program use was also high, with 62.80% (n=2509) of the invited mothers having visited the website. In addition, on average, they spent quite some time on the website (22 min). Taken together, the intervention has potential for broad national dissemination and implementation.

Furthermore, subgroup analysis with 3 intention groups (ie, negative, in doubt, and positive) showed that the intervention had the most positive effects on decisional conflict for mothers who were doubting. For mothers with a negative intention, the most positive effects were found on intention, attitude, and subjective norms. This is promising, as for a population-wide program it is relevant to guide those in doubt toward making an informed choice without decisional conflict and to persuade those having negative intentions toward vaccination. Fortunately, we did not find any adverse effects of the intervention in mothers with a negative intention, such as a further decrease in their intention or attitude.

Methodological Considerations

There are two methodological considerations. First, with large sample sizes, as in this study, even small effects can become statistically significant [54]. However, the positive intervention effects were consistently found on almost all outcomes. In addition, the large sample provided us with sufficient data for conducting analyses on subsamples (ie, based on sociodemographics, sample, and three intention groups) while maintaining sufficient levels of power [54]. The effect sizes that we found are in line with other Web-based interventions targeting health behavior outcomes [55]. We believe these, even small effects, are of relevance in public health as they become substantial at the population level. In addition, the intervention was of help for those in doubt and did not have any detrimental effect. We therefore find this Web-based tailored program is a substantial step forward in improving both research and practice in the context of the promotion of HPV vaccination acceptability.

Second, we used a scale comprising two items of the Self-Report Habit Index for measuring habit [41]. This accounted for the extent to which getting the HPV vaccination was something mothers did (1) naturally and (2) without thinking. The positive intervention effect on the composite measure might indicate that the intervention induced perceptions about the HPV vaccination as something you take for granted, without thinking. The latter is unwanted, considering the aim of initiating active processing of information about the vaccination. Fortunately, secondary analysis separating the two items showed that there was a positive effect of the intervention on the first item (ie, "naturally") but not on the second (ie, "without thinking"). In retrospect, the label "habit" attributed to the two-item operationalization appears misleading, though the intervention strengthened the mothers' belief about getting the HPV vaccination as something natural.

Strengths and Limitations

Important strengths of this study are the randomized controlled design, adequate sample sizes, and the reliable objective assessment of HPV vaccination uptake. Furthermore, the systematic, stepwise development of the intervention was a notable strength, as well as the mothers' positive subjective evaluation of the intervention and the objectively assessed high level of program use.

Some limitations of this study should be considered as well. First, the study was subject to a considerable amount of attrition. Unfortunately, attrition is quite common in studies on Web-based interventions [56]. Dropout analysis showed that dropout was selective. For instance, there were higher dropout rates in the intervention condition, which has also been reported for other (tailored) interventions [57-59]. In this study, we handled the missing data and selective dropout by using multiple imputation [60]. Results from the complete case analyses only slightly differed from those from the ITT analyses. Thus, it seems unlikely that the observed effects are spurious or due to selective dropout. Second, caution is needed when generalizing the results of this study to the general population (ie, Dutch mothers of girls aged 12 years) because we were unable to check the sample's representativeness. However, we did not find any differences in effectiveness of the intervention in specific subgroups of participants, as indicated by the conducted moderation analyses with sociodemographics.

Conclusions and Recommendations

The study findings suggest that this Web-based tailored intervention has the potential to improve both HPV vaccination acceptability and IDM, and decrease decisional conflict among mothers of invited girls. Therefore, we recommend nationwide dissemination and implementation. Furthermore, we recommend developing (tailored) interventions targeting the daughter and the mother's partner, as these have appeared to be important social referents [7,8]. Research has indicated that tailoring the intervention could have beneficial effects for girls, as they also expressed their need for interactive and personal information about the HPV vaccination [8]. This still needs to be investigated for the partners. In addition, boys may also become a relevant target group. In other countries, such as Australia, boys are already included in the national immunization program. This may contribute to the achievement of herd immunity and to a reduction of the global burden of a variety of HPV-related cancers in women and men [61-63].

Acknowledgments

We would like to thank Petra Oomen (RIVM) and the panels (Veldkamp BV, Intromart GFK, and NGO Flycatcher) for their help with data collection. ZonMw (The Netherlands Organization for Health Research and Development) and NWO (Netherlands Organization for Scientific Research) support this study (grant-number: 50-51515-98-259).

Conflicts of Interest

None declared.

Multimedia Appendix 1

Intervention screenshots.

[PDF File (Adobe PDF File), 618KB - jmir_v19i9e312_app1.pdf]

Multimedia Appendix 2

CONSORT-EHealth (V 1.6.1).

[PDF File (Adobe PDF File), 533KB - jmir_v19i9e312_app2.pdf]

References

1. Ferlay J, Shin H, Bray F, Forman D, Mathers C, Parkin DM. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. Int J Cancer 2010 Dec 15;127(12):2893-2917 [FREE Full text] [doi: 10.1002/ijc.25516] [Medline: 21351269]
2. Schiffman M, Castle PE. Human papillomavirus: epidemiology and public health. Arch Pathol Lab Med 2003 Aug;127(8):930-934. [doi: 10.1043/1543-2165(2003)127<930:HPEAPH>2.0.CO;2] [Medline: 12873163]
3. Bruni L, Barrionuevo-Rosas L, Albero G, Serrano B, Mena M, Gómez D, et al. Hpvcentre. ICO Information centre on HPV and cancer (HPV Information centre) URL: http://www.hpvcentre.net/statistics/reports/XWX.pdf [accessed 2017-04-19] [WebCite Cache ID 6ppyIMFdK]
4. Braspenning J, Tacken M, Penders A, Van DHH, De Bakker D. Nivel. 2001. Opkomst bevolkingsonderzoek baarmoederhalskanker, 1997-1999 URL: https://www.nivel.nl/node/2430?database=ChoicePublicat&priref=1459 [WebCite Cache ID 6ssrTVaWV]
5. Gezondheidsraad. Vara. 2008. Vaccinatie tegen baarmoederhalskanker URL: https://zembla.vara.nl/pdf/GR%20baarmoederhalskanker.pdf [WebCite Cache ID 6ssrl7JzZ]
6. van Lier E, Oomen P, Giesbers H, van Vliet JA, Drijfhout I, Zonnenberg-Hoff I, et al. Rivm. 2016. Vaccinatiegraad Rijksvaccinatieprogramma Nederland: verslagjaar URL: http://www.rivm.nl/Documenten_en_publicaties/Wetenschappelijk/Rapporten/2016/juni/Vaccinatiegraad_Rijksvaccinatieprogramma_Nederland_Verslagjaar_2016 [accessed 2017-01-31] [WebCite Cache ID 6nvO6BTOw]

7. van KH, Otten W, Ruiter RA, Fekkes M, van SJ, Dusseldorp E, et al. Determinants of HPV vaccination intentions among Dutch girls and their mothers: a cross-sectional study. BMC Public Health 2013 Feb 06;13:111 [FREE Full text] [doi: 10.1186/1471-2458-13-111] [Medline: 23388344]
8. Van Keulen M, Otten W, Ruiter R, Van Steenbergen J, Fekkes M, Paulussen TW. Researchgate. 2013. Redenen om zich te laten vaccineren tegen HPV URL: https://www.researchgate.net/profile/Jim_E_Steenbergen/publication/236327362_Reasons_for_having_oneself_vaccinated_against_HPV_Implications_for_the_future_provision_of_information/links/00b4951e1b971b37eb000000/Reasons-for-having-oneself-vaccinated-against-HPV-Implications-for-the-future-provision-of-information.pdf?origin=publication_list [WebCite Cache ID 6sss8ydeu]
9. Bartholomew EL, Markham C, Ruiter R, Fernàndez ME, Kok G, Parcel GS. Planning health promotion programs: An Intervention Mapping approach, 4th edition. Hoboken, NJ: Wiley; 2016.
10. Bennett AT, Patel DA, Carlos RC, Zochowski MK, Pennewell SM, Chi AM, et al. Human papillomavirus vaccine uptake after a tailored, online educational intervention for female university students: a randomized controlled trial. J Womens Health (Larchmt) 2015 Nov;24(11):950-957 [FREE Full text] [doi: 10.1089/jwh.2015.5251] [Medline: 26488269]
11. Hopfer S. Effects of a narrative HPV vaccination intervention aimed at reaching college women: a randomized controlled trial. Prev Sci 2012 Apr;13(2):173-182. [doi: 10.1007/s11121-011-0254-1] [Medline: 21993613]
12. Gerend MA, Shepherd MA, Lustria ML. Increasing human papillomavirus vaccine acceptability by tailoring messages to young adult women's perceived barriers. Sex Transm Dis 2013 May;40(5):401-405. [doi: 10.1097/OLQ.0b013e318283c8a8] [Medline: 23588130]
13. Grandahl M, Rosenblad A, Stenhammar C, Tydén T, Westerling R, Larsson M, et al. School-based intervention for the prevention of HPV among adolescents: a cluster randomised controlled study. BMJ Open 2016 Jan 27;6(1):e009875 [FREE Full text] [doi: 10.1136/bmjopen-2015-009875] [Medline: 26817639]
14. Neville LM, O'Hara B, Milat AJ. Computer-tailored dietary behaviour change interventions: a systematic review. Health Educ Res 2009 Aug;24(4):699-720 [FREE Full text] [doi: 10.1093/her/cyp006] [Medline: 19286893]
15. Pot M, Van Keulen H, Ruiter RA, Eekhout I, Mollema L, Paulussen TW. Motivational and contextual determinants of HPV-vaccination uptake: a longitudinal study among mothers of girls invited for the HPV-vaccination. Prev Med 2017 Apr 04;100:41-49 [FREE Full text] [doi: 10.1016/j.ypmed.2017.04.005] [Medline: 28389328]
16. Marteau TM, Dormandy E, Michie S. A measure of informed choice. Health Expect 2001 Jun;4(2):99-108. [Medline: 11359540]
17. O'Connor AM, Jacobsen MJ, Stacey D. An evidence-based approach to managing women's decisional conflict. J Obstet Gynecol Neonatal Nurs 2002;31(5):570-581. [Medline: 12353737]
18. Ernst E, Pittler MH. Efficacy or effectiveness? J Intern Med 2006 Nov;260(5):488-490 [FREE Full text] [doi: 10.1111/j.1365-2796.2006.01707.x] [Medline: 17040256]
19. Kreuter MW, Wray RJ. Tailored and targeted health communication: strategies for enhancing information relevance. Am J Health Behav 2003 Dec;27 Suppl 3:S227-S232. [Medline: 14672383]
20. 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]
21. 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]
22. Prochaska J, Wright J, Velicer W. Evaluating theories of health behavior change: a hierarchy of criteria applied to the transtheoretical model. Appl Psychol 2008;57(4):561-588. [doi: 10.1111/j.1464-0597.2008.00345.x]
23. Fishbein M, Ajzen I. Predicting and changing behavior: The Reasoned Action Approach. New York: Taylor & Francis; 2010.
24. Petty R, Barden J, Wheeler S. The Elaboration Likelihood Model of Pursuasion: Developing health promotions for sustained behavioral change. In Diclimente RJ, Crosby RA. & Kegler M (Eds.), Emerging theories in health promotion practice and research, 2nd edition. San Francisco, CA: Jossey-Bass; 2009:185-214.
25. Miller W, Rollnick S. Motivational interviewing: Preparing people for change, 2nd edition. New York: The Guilford Press; 2002.
26. Whittingham J, Ruiter RA, Zimbile F, Kok G. Experimental pretesting of public health campaigns: a case study. J Health Commun 2008;13(3):216-229. [doi: 10.1080/10810730701854045] [Medline: 18569355]
27. Whittingham JR, Ruiter RA, Castermans D, Huiberts A, Kok G. Designing effective health education materials: experimental pre-testing of a theory-based brochure to increase knowledge. Health Educ Res 2008 Jun;23(3):414-426. [doi: 10.1093/her/cym018] [Medline: 17586587]
28. Beun R, de Vos E, Witteman E. Embodied conversational agentsffects on memory performance and anthropomorphisation. In: Proceedings of the International Conference on Intelligent Virtual Agents. New York: Springer-Verlag; 2003:315-319.
29. Gerend MA, Shepherd MA, Lustria MLA. Increasing human papillomavirus vaccine acceptability by tailoring messages to young adult women's perceived barriers. Sex Transm Dis 2013 May;40(5):401-405. [doi: 10.1097/OLQ.0b013e318283c8a8] [Medline: 23588130]

30. Baylor A, Kim Y. Pedagogical agent design: the impact of agent realism, gender, ethnicity, and instructional role. In: LNCS, intelligent tutoring systems, volume 3320. Berlin/Heidelberg: Germany Springer; 2004:592-603.
31. Jin SA. The effects of incorporating a virtual agent in a computer-aided test designed for stress management education: the mediating role of enjoyment. Comput Human Behav 2010 May;26(3):443-451. [doi: 10.1016/j.chb.2009.12.003]
32. Blanson HO, Rogers W, Fisk AD, Neerincx M, Lindenberg J, van der Mast CA. Usability of an adaptive computer assistant that improves self-care and health literacy of older adults. Methods Inf Med 2008;47(1):82-88 [FREE Full text] [Medline: 18213433]
33. Blanson HOA, van der Boog PJ, Lindenberg J, van der Mast CA, Neerincx MA, Zwetsloot-Schonk BJ. An online lifestyle diary with a persuasive computer assistant providing feedback on self-management. Technol Health Care 2009;17(3):253-267. [doi: 10.3233/THC-2009-0545] [Medline: 19641261]
34. Van der Pal S, Otten W, Detmar S. Evaluatie van de voorlichting aan ouders over de hielprik. TSG 2010;88:449-453. [doi: 10.1007/s12508-010-0679-7]
35. Altman DG, Royston P. The cost of dichotomising continuous variables. BMJ 2006 May 6;332(7549):1080 [FREE Full text] [doi: 10.1136/bmj.332.7549.1080] [Medline: 16675816]
36. Michie S, Dormandy E, Marteau TM. The multi-dimensional measure of informed choice: a validation study. Patient Educ Couns 2002 Sep;48(1):87-91. [Medline: 12220754]
37. Rubin D. Multiple Imputation for Nonresponse in Surveys. New York: Wiley; 1987.
38. Paulussen TG, Lanting CI, Buijs GB, Hirasing RA. Google. 2000. Ouders over het Rijksvaccinatieprogramma: Tevredenheid en vaccinatiebereidheid van ouders van jonge kinderen in Nederland URL: http://bit.ly/2wEgIuL [WebCite Cache ID 6ssvWVfBU]
39. Reiter PL, Brewer NT, Gottlieb SL, McRee A, Smith JS. Parents' health beliefs and HPV vaccination of their adolescent daughters. Soc Sci Med 2009 Aug;69(3):475-480. [doi: 10.1016/j.socscimed.2009.05.024] [Medline: 19540642]
40. Gerend MA, Weibley E, Bland H. Parental response to human papillomavirus vaccine availability: uptake and intentions. J Adolesc Health 2009 Nov;45(5):528-531. [doi: 10.1016/j.jadohealth.2009.02.006] [Medline: 19837361]
41. Verplanken B, Orbell S. Reflections on past behavior: a self-report index of habit strength. J Appl Social Pyschol 2003 Jun;33(6):1313-1330. [doi: 10.1111/j.1559-1816.2003.tb01951.x]
42. Van Buuren S. Flexible Imputation of Missing Data. Boca Raton: Chapman Hall/CRC; 2012.
43. White IR, Royston P, Wood AM. Multiple imputation using chained equations: issues and guidance for practice. Stat Med 2011 Feb 20;30(4):377-399. [doi: 10.1002/sim.4067] [Medline: 21225900]
44. Van Breukelen GJ. ANCOVA versus change from baseline: more power in randomized studies, more bias in nonrandomized studies [corrected]. J Clin Epidemiol 2006 Sep;59(9):920-925. [doi: 10.1016/j.jclinepi.2006.02.007] [Medline: 16895814]
45. R Development Core Team. R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing; 2009.
46. Cohen J. Statistical power analysis for the behavioral science. Hillsdale, NJ: Lawrence Erlbaum Associates Publishers; 1988.
47. Fayers PM, King MT. How to guarantee finding a statistically significant difference: the use and abuse of subgroup analyses. Qual Life Res 2009 Jun;18(5):527-530. [doi: 10.1007/s11136-009-9473-3] [Medline: 19343540]
48. Kata A. Anti-vaccine activists, Web 2.0, and the postmodern paradigm--an overview of tactics and tropes used online by the anti-vaccination movement. Vaccine 2012 May 28;30(25):3778-3789. [doi: 10.1016/j.vaccine.2011.11.112] [Medline: 22172504]
49. McGuire W. In: Berkowitz L, editor. Advances in experimental social psychology, volume I. New York: Academic Press; 1964.
50. Paulussen TG, Hoekstra F, Lanting CI, Buijs GB, Hirasing RA. Determinants of Dutch parents' decisions to vaccinate their child. Vaccine 2006 Jan 30;24(5):644-651. [doi: 10.1016/j.vaccine.2005.08.053] [Medline: 16157423]
51. Integraal Kankercentrum Nederland. Cijfersoverkanker. 2016. Cijfers over kanker URL: http://www.cijfersoverkanker.nl/selecties/Dataset_1/img58906e45e92b1 [accessed 2017-01-31] [WebCite Cache ID 6nvPAOOSA]
52. Ruiter RA, Abraham C, Kok G. Scary warnings and rational precautions: a review of the psychology of fear appeals. Psychol Health 2001 Nov;16(6):613-630. [doi: 10.1080/08870440108405863]
53. Hofman R, van Empelen P, Richardus JH, de Kok IM, de Koning HJ, van Ballegooijen M, et al. Predictors of HPV vaccination uptake: a longitudinal study among parents. Health Educ Res 2014 Feb;29(1):83-96. [doi: 10.1093/her/cyt092] [Medline: 24041721]
54. Lin M, Lucas HC, Shmueli G. Too big to fail: large samples and the p-value problem. Inf Syst Res 2013 Dec;24(4):906-917. [doi: 10.1287/isre.2013.0480]
55. Webb TL, Joseph J, Yardley L, Michie S. Using the internet to promote health behavior change: a systematic review and meta-analysis of the impact of theoretical basis, use of behavior change techniques, and mode of delivery on efficacy. J Med Internet Res 2010;12(1):e4 [FREE Full text] [doi: 10.2196/jmir.1376] [Medline: 20164043]
56. Kohl LFM, Crutzen R, de Vries NK. Online prevention aimed at lifestyle behaviors: a systematic review of reviews. J Med Internet Res 2013;15(7):e146 [FREE Full text] [doi: 10.2196/jmir.2665] [Medline: 23859884]

57. Oenema A, Brug J, Dijkstra A, de WI, de VH. Efficacy and use of an internet-delivered computer-tailored lifestyle intervention, targeting saturated fat intake, physical activity and smoking cessation: a randomized controlled trial. Ann Behav Med 2008 Apr;35(2):125-135. [doi: 10.1007/s12160-008-9023-1] [Medline: 18363076]
58. Oenema A, Tan F, Brug J. Short-term efficacy of a web-based computer-tailored nutrition intervention: main effects and mediators. Ann Behav Med 2005 Feb;29(1):54-63. [doi: 10.1207/s15324796abm2901_8] [Medline: 15677301]
59. Spittaels H, De Bourdeaudhuij BI, Vandelanotte C. Evaluation of a website-delivered computer-tailored intervention for increasing physical activity in the general population. Prev Med 2007 Mar;44(3):209-217. [doi: 10.1016/j.ypmed.2006.11.010] [Medline: 17197015]
60. Twisk J. Applied Multilevel Analysis. Cambridge: Cambridge University Press; 2006.
61. Georgousakis M, Jayasinghe S, Brotherton J, Gilroy N, Chiu C, Macartney K. Population-wide vaccination against human papillomavirus in adolescent boys: Australia as a case study. Lancet Infect Dis 2012 Aug;12(8):627-634. [doi: 10.1016/S1473-3099(12)70031-2]
62. Wakeham K, Kavanagh K. The burden of HPV-associated anogenital cancers. Curr Oncol Rep 2014 Sep;16(9):402. [doi: 10.1007/s11912-014-0402-4] [Medline: 25118645]
63. Braakhuis BJM, Leemans CR, Visser O. Incidence and survival trends of head and neck squamous cell carcinoma in the Netherlands between 1989 and 2011. Oral Oncol 2014 Jul;50(7):670-675 [FREE Full text] [doi: 10.1016/j.oraloncology.2014.03.008] [Medline: 24735546]

Abbreviations

HPV: human papillomavirus
IDM: informed decision making
ITT: intention-to-treat analysis
NIP: National Immunization Program
RCT: randomized controlled trial
SD: standard deviation

Edited by J Torous, T Kool; submitted 06.02.17; peer-reviewed by M Grandahl, J Glanz, N Bragazzi; comments to author 26.03.17; revised version received 09.05.17; accepted 03.06.17; published 06.09.17.

Please cite as:
Pot M, Paulussen TGWM, Ruiter RAC, Eekhout I, de Melker HE, Spoelstra MEA, van Keulen HM
Effectiveness of a Web-Based Tailored Intervention With Virtual Assistants Promoting the Acceptability of HPV Vaccination Among Mothers of Invited Girls: Randomized Controlled Trial
J Med Internet Res 2017;19(9):e312
URL: http://www.jmir.org/2017/9/e312/
doi:10.2196/jmir.7449
PMID:28877862

Original Paper

User Acceptance of Computerized Cognitive Behavioral Therapy for Depression: Systematic Review

Theresia Rost[1], MA; Janine Stein[1], Dr rer med; Margrit Löbner[1], Dr rer med; Anette Kersting[2], Prof Dr med; Claudia Luck-Sikorski[3,4], Prof, Dr rer med; Steffi G Riedel-Heller[1], MPH, Prof Dr med

[1]Institute of Social Medicine, Occupational Health and Public Health, Medical Faculty, University of Leipzig, Leipzig, Germany
[2]Department of Psychosomatic Medicine, University Medical Centre, University of Leipzig, Leipzig, Germany
[3]Mental Health and Psychotherapy, Stiftung Rehabilitation Heidelberg University of Applied Health Services, Gera, Germany
[4]Integrated Research and Treatment Center (IFB) AdiposityDiseases, University Hospital, Leipzig, Germany

Corresponding Author:
Steffi G Riedel-Heller, MPH, Prof Dr med
Institute of Social Medicine, Occupational Health and Public Health
Medical Faculty
University of Leipzig
Building A, 3rd Fl.
Philipp-Rosenthal-Str. 55
Leipzig, 04103
Germany
Phone: 49 3419715408
Fax: 49 3419724569
Email: Steffi.Riedel-Heller@medizin.uni-leipzig.de

Abstract

Background: Computerized cognitive behavioral therapy (cCBT) has been proven to be effective in depression care. Moreover, cCBT packages are becoming increasingly popular. A central aspect concerning the take-up and success of any treatment is its user acceptance.

Objective: The aim of this study was to update and expand on earlier work on user acceptance of cCBT for depression.

Methods: This paper systematically reviewed quantitative and qualitative studies regarding the user acceptance of cCBT for depression. The initial search was conducted in January 2016 and involved the following databases: Web of Science, PubMed, the Cochrane Library, and PsycINFO. Studies were retained if they described the explicit examination of the user acceptance, experiences, or satisfaction related to a cCBT intervention, if they reported depression as a primary outcome, and if they were published in German or English from July 2007 onward.

Results: A total of 1736 studies were identified, of which 29 studies were eligible for review. User acceptance was operationalized and analyzed very heterogeneously. Eight studies reported a very high level of acceptance, 17 indicated a high level of acceptance, and one study showed a moderate level of acceptance. Two qualitative studies considered the positive and negative aspects concerning the user acceptance of cCBT. However, a substantial proportion of reviewed studies revealed several methodical shortcomings.

Conclusions: In general, people experience cCBT for depression as predominantly positive, which supports the potential role of these innovative treatments. However, methodological challenges do exist in terms of defining user acceptance, clear operationalization of concepts, and measurement.

(J Med Internet Res 2017;19(9):e309) doi:10.2196/jmir.7662

KEYWORDS

computerized cognitive behavior therapy; depression; user acceptance; systematic review; review; behavior therapy; depressive disorder

Introduction

Depressive disorders are among the most common and serious mental illnesses [1]. Globally, 350 million people of all ages are estimated to suffer from depression. If depressive disorders are detected at an early stage, they are highly treatable in the majority of cases [2]. There are known effective psychological treatments, for example, cognitive behavioral therapy (CBT) [3]. However, individuals suffering from depression often find themselves confronted with barriers to receiving appropriate care such as social stigma associated with mental disorders, long waiting times, or the logistical difficulties of appearing in person for treatment [4,5]. For these reasons, computerized programs present an innovative approach to improving access to psychological treatments for depression. There is evidence that computerized cognitive behavioral therapy (cCBT) is effective in the treatment of various mental disorders, including depression [6-10]. There are a number of advantages that are associated with cCBT such as anonymity, wide availability, or location-independent and around-the-clock access [9,11]. Well-known cCBT programs such as Beating The Blues and MoodGYM have been shown to provide a promising option for the treatment of mental health problems [9,12,13]. A prerequisite for cCBT programs to be effective is its user acceptance, as the implementation of an innovative intervention such as cCBT can be affected negatively because of individuals being unwilling to use it. For example, the absence of a contact person and the resulting anonymity can have a negative impact on the user's motivation to start or keep up with a cCBT program. Therefore, it is of utmost importance to consider user acceptance when developing and implementing a cCBT program for the treatment of depression.

The concept of user acceptance arose as a key term in the scientific discourse. Definitions of the term differ widely depending on the intended use [14]. One of the most popular approaches is the technology acceptance model (TAM) developed by Davis [15]. TAM illustrates user acceptance determined by two factors: perceived usefulness and perceived ease of use. According to Davis [15], both have a significant impact on a person's attitude toward using a new technology. Kollmann [16] and Rogers [17] went one step further and combined different phases in their acceptance models. Therefore, the user passes through phases from getting to know a new technology, to forming an attitude toward it, to a decision whether to use or not to the confirmation of the decision. On this basis, user acceptance can be defined as the willingness of individuals to employ information technology for the tasks it is designed to support, the realization, and approval of the decision to employ. All of these models have one thing in common: user acceptance is considered to be a process beginning with an attitude toward the innovation and developing into satisfaction with the innovation; it is not an instantaneous act. Accordingly, we have conceived acceptance as the act of accepting, experiencing, and being satisfied.

Since the emergence of the first cCBT programs, there have been a number of reviews addressing the user acceptance of cCBT; however, they have utilized different approaches. In their reviews, Titov [18], Andrews et al [8], and Vallury et al [19] focused broadly on effectiveness and user acceptance of cCBT for several mental disorders, including depression and anxiety disorders. Waller and Gilbody [20] reviewed quantitative and qualitative studies examining adverse consequences, accessibility, and acceptability of cCBT programs for treating anxiety and depression. However, Kaltenthaler et al [21] provide the only review with a very comprehensive and focused insight into the user acceptance of cCBT for depression, including research up to June 2007. They systematically reviewed sources of information on acceptability to patients of cCBT for depression. As a result, they documented several studies reporting positive expectancies and high satisfaction in routine care cCBT services for those completing the treatment and argued that studies should reveal more detailed information on patient recruitment methods, dropout rates, and reasons for dropping out. Furthermore, they drafted well-designed surveys and qualitative studies included alongside trials to determine levels of patient acceptability as implications for further research.

On this basis, we provide a systematic overview on user acceptance of cCBT for depression over the last 10 years and widen the perspective to include the notion that the process of user acceptance spans a number of phases, including accepting, experiencing, and being satisfied with cCBT. We intend to answer the following research questions: (1) which measures were used to examine the user acceptance of cCBT for depression? and (2) to what degree do users accept cCBT for depression?

Methods

This systematic review was conducted according to guidelines from the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement [22]. On the basis of the PICO (Patient, problem, or population; Intervention; Comparison, control, or comparator; Outcome) approach and the review by Kaltenthaler et al [21], the criteria for inclusion were as follows.

Eligibility Criteria

Population

Since cCBT programs may be also designed for people not undergoing medical treatment, we decided to widen the focus on people with or without medical attention. Thus, studies with participants of all ages with a diagnosis of depression of all degrees of severity were regarded as eligible for inclusion in this review.

Interventions

All cCBT interventions and their subtypes (eg, mindfulness-based cognitive therapy and behavioral activation) delivered alone or as part of a package of care via the Internet were taken into consideration.

Comparison

Randomized controlled trials (RCTs), nonrandomized comparative trials, noncomparative trials, and qualitative studies published from July 2007 to January 2016 were included.

Outcome

Studies were included if they reported on the following: data on user acceptance in terms of acceptability, satisfaction, or experiences concerning cCBT; studies with depression as a primary outcome; and studies providing information on study design and measures, including a description of the delivered treatment and the sample including the number, age, and sex of participants. Studies were excluded if they were not reported in English or German or if they were single case reports.

Search

The search for relevant literature was conducted in four bibliographic databases from July 1, 2007 to January 31, 2016, which are as follows: Web of Science, PubMed, the Cochrane Library, and PsycINFO. Furthermore, the bibliographies of identified papers were searched to identify other potentially eligible papers. Since studies about user acceptance emanate from a young research area, it is conceivable that many studies measure or report about it via proxy indices, which was considered in the search strategy of this study. Considering British and American spelling, a search strategy combining the following search terms was used to ensure complete coverage of studies: Concept 1 (“internet” OR “web” OR “DVD” OR “CD-ROM” OR “online” OR “computer*” OR “e-health” OR “electronic” OR “program” OR “programme”) AND Concept 2 (“CCBT” OR “CBT” OR “cognitive therapy” OR “behavior therapy” OR “behavioral therapy” OR “behavioural therapy” OR “behaviour therapy”) AND Concept 3 (“accept*” OR “satisfaction” OR “adherence” OR “compliance” OR “take up rates” OR “patient dropout rates” OR “reasons for dropout” OR “patient drop-out rates” OR “reasons for drop-out”) AND Concept 4 (“depress*” OR “dysthym*” OR “mood disorder” OR “affective disorder” OR “melancholia”).

Study Selection

After removing duplicates identified in databases and reference lists, titles and abstracts of the texts were scanned to examine indications for meeting the inclusion criteria. For all remaining papers that deemed relevant, the full text was reviewed. All information from the included studies was gathered by one reviewer and checked by a second.

Data Collection Process

We extracted information on the characteristics of the program, as well as information on the study design, the setting, the ways of recruitment, the sample, dropout and completion rates, and, if available, reasons for dropout.

Synthesis of Results

To allow a better comparability, we transferred the results into levels of acceptance that range from low (−−) to moderate (−) to high (+) to very high (++). The levels follow the results reported in percentage and scale values that were assigned to quartiles. Therefore, results ranging between 0% and 25% were assigned to *low*, between 26% and 50% to *moderate*, 51% and 75% to *high*, and 76% and 100% to *very high*. This also applies equivalently to scale values. For example, Danaher et al [23] used a 4-point scale for the elicitation of satisfaction (1=not at all satisfied, 4=very satisfied), for which the quartiles are as follows: 1 to 1.75 (low), 1.76 to 2.5 (moderate), 2.6 to 3.25 (high), and 3.26 to 4 (very high). The study reported mean scores of 3.3 (satisfaction with features of the program) and 3.4 (helpfulness of personal coach calls). Thus, the level of acceptance can be described as *very high*.

If there were considerations of positive and negatives aspects concerning the user acceptance of cCBT, they were characterized (~).

Results

Study Selection

As shown in Figure 1, a total of 1736 potentially relevant papers were identified through database searching; 36 additional papers were identified from reference lists. After removing a total of 564 duplicates, 1208 papers were screened based on their titles and abstracts. A total of 1123 publications were excluded because they did not fulfill the inclusion criteria. Hence, the remaining 85 full-text papers were assessed for eligibility. Of those, 56 publications were excluded mainly because depression was not defined as a primary outcome (n=28), and user acceptance was not examined as described in their titles or abstracts (n=17). Furthermore, one study could not be taken into consideration because of a highly selected sample comprising caregivers of anorexia nervosa patients. Finally, we included 29 studies for further analysis.

Study Characteristics

Objects of investigation were several cCBT programs, including “MoodGYM” and “Beating the Blues” that were examined most commonly. Table 1 presents further information on the characteristics of the cCBT programs. As Multimedia Appendix 1 shows, 16 of included studies were RCTs, 8 were noncomparative trials, 3 were qualitative studies, and 2 were comparative but nonrandomized trials. Five of the studies had a special feature: three studies made a comparison between guided and unguided programs [24-26] and the remaining two studies compared two programs [27,28].

Figure 1. Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flowchart of the study selection and eligibility process.

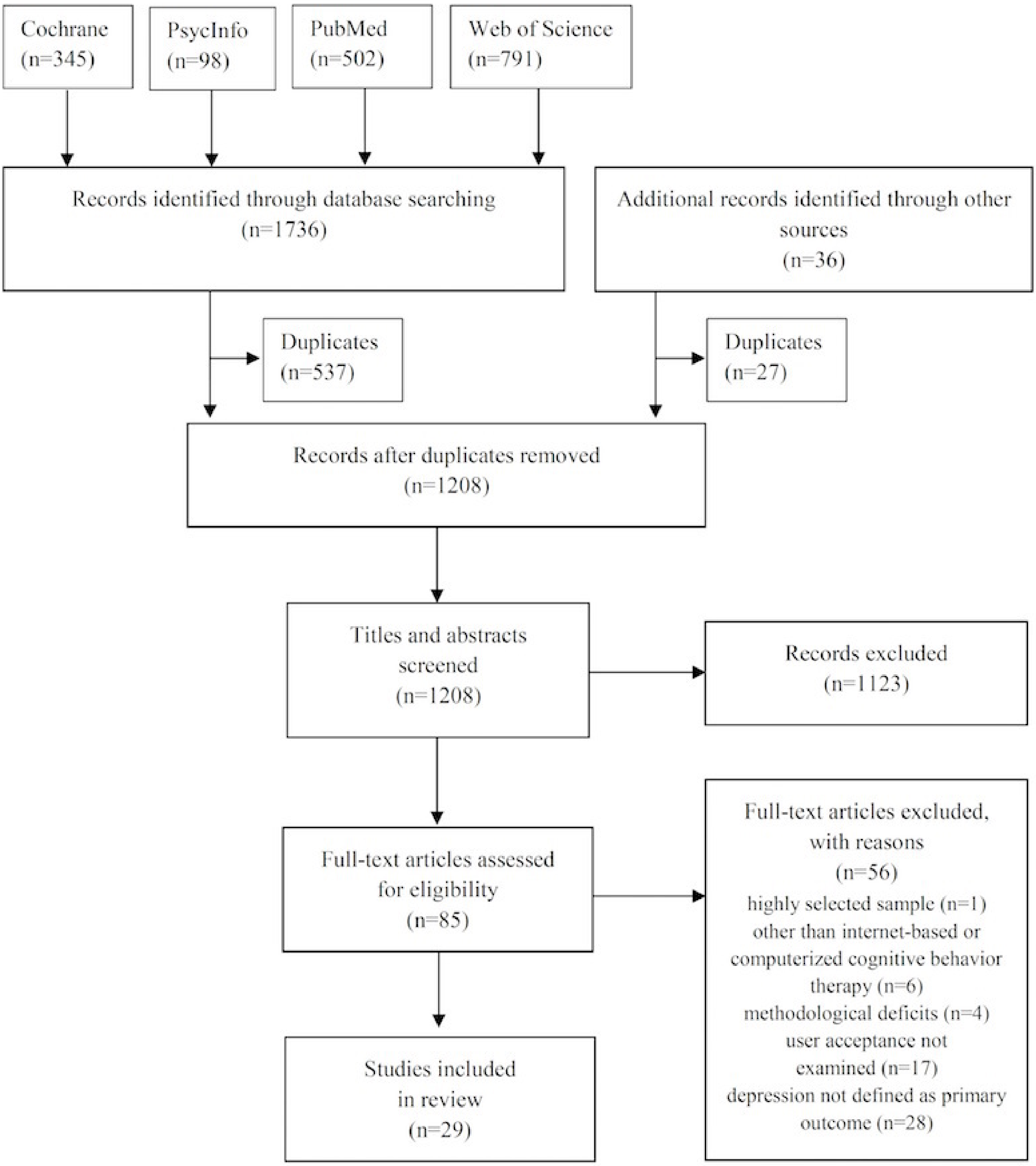


Table 1. Characteristics of the computerized cognitive behavioral therapy (cCBT) programs.

Author, year, country	Description of the program	Support (nontherapeutic support; therapist support)
Ahmedani et al, 2015, United States [29]	iCBT[a] program (no name), brief tailored mobile health intervention, based on a combination of motivational interviewing and CBT[b] models	yes; no
Berger et al, 2011, Switzerland [24]	Deprexis, a self-help program comprising 10 content modules and a summary module covering a variety of therapeutic content that is broadly consistent with a cognitive behavioral model IG[c] I: unguided self-help condition IG II: guided self-help condition	no; yes (participants randomized in guided self-help condition received email contact with a therapist)
Berman et al, 2014, United States [30]	ePST, a 6-session, stand-alone multimedia, interactive, computer-based problem-solving treatment	yes; no
Boeschoten et al, 2012, The Netherlands [31]	cCBT[d] program (no name), based on the original "problem-solving therapy," adjusted for multiple sclerosis patients with comorbid depression and comprising 5 modules containing text, exercise, and examples	yes; no
Burns et al, 2011, United States [32]	Mobilyze!, an 8-week mobile phone– and Internet-based intervention for depression	yes; yes
Cartreine et al, 2012, United States [33]	ePST, a 6-session, stand-alone multimedia, interactive, computer-based problem-solving treatment	not reported
Choi et al, 2012, Australia [34]	The Brighten Your Mood Program, a cultural adapted version of the Sadness Program comprising 6 lessons	yes; no
Danaher et al, 2013, United States [23]	MomMoodBooster , an interactive guided Web-based intervention for postpartum depression comprising 6 sessions	yes
de Graaf et al, 2009, The Netherlands [35]	Colour Your Life, a Web-based multimedia, interactive, self-help cCBT program for depression based on the Dutch version "Coping With Depression course" and comprises 8 weekly sessions Colour Your Life + TAU[e]	no; only participants who got the intervention + TAU
Dear et al, 2013, Australia [36]	Managing Your Mood, a structured 5-lesson Web-based intervention that encourages participants to learn and practice core CBT psychological skills	yes
Dimidjian et al, 2014, United States [37]	Mindful Mood Balance, a Web-based, 8-session self-administered platform	not reported
Geraedts et al, 2015, The Netherlands [38]	Happy@Work, a brief 6-lesson Internet intervention based on problem-solving treatment, cognitive therapy, and a guideline for employees to help them prevent work-related stress	yes
Gerhards et al, 2011, The Netherlands [39]	Colour Your Life , a Web-based multimedia, interactive computer program for depression comprising 8 weekly sessions and a 9th booster session	no; no
Hind et al, 2010, United Kingdom [27]	Beating the Blues, an interactive computer program with 8 modules for the treatment of depressive and anxiety disorders MoodGYM, a freeware cCBT program comprising 5 modules IG I: Beating the Blues IG II: MoodGYM	yes
Høifødt et al, 2013, Norway [12]	MoodGYM, a Web-based program containing 5 modules comprising written information, animations, interactive exercises, and quizzes	no; yes
Kay-Lambkin et al, 2011, Australia [40]	SHADE, a clinician-assisted computer-based psychological treatment comprising 10 sessions and delivered on DVD	yes
Knowles et al, 2015, United Kingdom [28]	MoodGYM, a Web-based program containing 5 modules comprising written information, animations, interactive exercises, and quizzes Beating the Blues, an interactive computer program with 8 modules for the treatment of depressive and anxiety disorders	no
Kok et al, 2014, The Netherlands [41]	Depressionfree, comprising Internet-based preventive cognitive therapy with 8 modules, telephone-delivered psychotherapy and mood monitoring Depressionfree + TAU	yes; yes

Author, year, country	Description of the program	Support (nontherapeutic support; therapist support)
Lintvedt et al, 2013, Norway [42]	MoodGYM, a self-help program based on principles of CBT, interpersonal therapy, and relaxation techniques comprising 5 modules BluePages provides evidence-based information about depression	no; no
Lucassen et al, 2014, New Zealand [43]	Rainbow SPARX, an interactive fantasy game comprising 7 modules designed to deliver CBT for the treatment of clinically significant depression; customized for sexual minority youth	yes
McMurchie et al, 2013, United Kingdom [44]	Beating the Blues, an interactive computer program with 8 modules for the treatment of depressive and anxiety disorders Beating the Blues + TAU	yes; no
Merry et al, 2012, New Zealand [45]	SPARX, an interactive fantasy game comprising 7 modules designed to deliver CBT for the treatment of clinically significant depression	no; no
O'Mahen et al, 2013, United Kingdom [46]	Postnatal Internet-based behavioral activation (iBA[f]), adapted for postnatal Web-based delivery from the manual developed for behavioral activation, comprising 11 weekly sessions Postnatal iBA + TAU	Access to Netmums' general depression chat room monitored by parent supporters and specialist health visitors
Perini et al, 2009, Australia [47]	The Sadness Program, a cCBT comprising 6 Web-based lessons, homework assignments, participation in an online discussion forum, and regular email contact with a mental health clinician	yes; yes
Richards and Timulak, 2013, Ireland [25]	Beating the Blues, an interactive computer program with 8 modules for the treatment of depressive and anxiety disorders IG I: Self-administered Beating the Blues IG II: Therapist-delivered Beating the Blues	no; only participants who were treated with intervention II got additional support from a therapist
Schneider et al, 2014, United Kingdom [48]	MoodGYM, a Web-based program containing 5 modules comprising written information, animations, interactive exercises, and quizzes	yes; no
Sheeber et al, 2012, United States [49]	Mom-Net program, an 8-session, Internet-facilitated CBT treatment for subthreshold and full syndrome depression, tailored to mothers of young children; the content foundation for the program was the Coping With Depression course IG I: Internet-facilitated intervention IG II: Delayed intervention or facilitated TAU	yes
Stasiak et al, 2014, New Zealand [50]	The journey, a cCBT with 7 modules of well-established core cognitive behavioral therapy techniques.	yes; no
Titov et al, 2010, Australia [26]	The Sadness Program, a cCBT program comprising 6 Web-based lessons, printable summary and homework assignments, automatic emails, and additional resource documents IG I: Technician-assisted group IG II: Clinician-assisted group	yes; yes

[a] iCBT: Internet-based cognitive behavioral therapy.

[b] CBT: cognitive behavioral therapy.

[c] IG: intervention group.

[d] cCBT: computerized cognitive behavioral therapy.

[e] TAU: treatment-as-usual.

[f] iBA: Internet-based behavioral activation.

Measures of Acceptance

As illustrated in Multimedia Appendix 2, the studies made use of several measures to examine the user acceptance of cCBT. The large majority of studies (n=25) used direct measures such as questionnaires or qualitative methods; two studies used indirect measures such as take-up rates, completion rates, or dropout rates; and two studies used a combination of direct and indirect measures. Of those using direct measures, four employed qualitative methods, five made use of well-established questionnaires, and 16 used study-specific questionnaires. These study-specific questionnaires varied substantially in their level of complexity. For example, Dear et al [36] ascertained the user acceptance of cCBT through 2 questions: (1) would you recommend the program to a friend? and (2) was the program worth your time? Berman et al [30] and Cartreine et al [33] employed the Acceptability of Self-Guided Treatment Questionnaire (AST) with 16 statements to be rated on a 7-point scale. The majority of study-specific developed questionnaires were left unspecified or roughly outlined (see Multimedia Appendix 2).

There were four studies that ascertained take-up, dropout, or completion rates as a means of assessing the user acceptance of cCBT. As Multimedia Appendix 3 shows, plenty of studies provided dropout and completion rates by default. Therefore, 17 studies revealed information about program completion, three reported dropout rates, seven trials commented on both, and two studies did not give any information about completion or dropout rates. For those reporting on rates for completing the entire program, the mean percentage of completion was 67.17% (standard deviation [SD] 20.29) with a range of 26.7% to 100%. With regard to the trials that compared guided with unguided programs, highly varying completion rates have been reported. Whereas Berger et al [24] documented that 36% of participants in the unguided self-help condition and 56% of participants in the guided self-help condition completed the entire program, Richards and Timulak [25] reported a completion rate of 16.28% in the unguided condition and 8.11% in the guided condition.

For the studies reporting on dropout rates, the mean percentage of dropout rates was 31.5% (SD 19.49), with a range of 0% to 63%. Twelve trials listed reasons for dropout. The most commonly stated reasons were a lack of time (n=6), technical difficulties, or computer-related problems (n=4), or participants experiencing the treatment as inconvenient (n=4). Since the trials differed in terms of study design, the extent of disclosure, and definitions of dropout and completion, it was difficult to draw comparisons between them regarding completion and dropout rates. Moreover, four studies documented take-up rates as follows: 83.3% [39], 56.9% [44], 97% [40], and 39% [12].

User Acceptance: Acceptability, Satisfaction, and Experiences

Multimedia Appendix 2 shows the results of all eligible studies regarding the user acceptance in terms of acceptability, satisfaction, and experiences with cCBT for depression. As shown here, results were primarily reported descriptively (eg, by reference to rated statements or responses to questions in a qualitative dimension) irrespective of whether the study type was quantitative or qualitative. According to the levels adapted by the author, eight studies (28%) reported a very high level of acceptance [23,26,36,43,45,47-49], 17 studies (59%) indicated a high level of acceptance [12,24,25,29-35,38,40-42,44,46,50], and one study (3%) showed a moderate level of acceptance [27]. No study showed a low level of user acceptance. One study (3%) gave conflicting information [37], which is why an allocation to a level of acceptance was not possible.

Two qualitative studies (7%) referred to considerations of positive and negative aspects concerning the user acceptance of cCBT. Specifically, Gerhards et al [39] and Knowles et al [28] focused on differentiated perceptions of cCBT that they extracted from qualitative interviews with participants. Gerhards et al [39] described the main barriers to be a lack of identification with cCBT, an absence of support to adhere to the program to gain deeper understanding, and inadequate computer and Internet skills. Motivators included the opportunity to use the program independent of time of day and location and added support as an improvement with regard to adherence and the course content. Knowles et al [28] showed that the same aspects of cCBT could be perceived positively and negatively, depending on the participant's experience and preference. For example, anonymity was associated with reduced pressure as compared with being face-to-face; however, it was also experienced as isolating and enhancing the feeling of loneliness. Similarly, flexibility was experienced as positive because patients are afforded a high degree of control but also as negative because the program can be seen as to be easy to avoid and difficult to sustain.

One special feature of three of the studies is the comparison of guided and unguided programs [24-26]. Whereas Titov et al [26] did not ascertain any differences concerning the satisfaction between the clinician-assisted cCBT and the technician-assisted cCBT, Berger et al [24] found evidence that participants in the guided condition were a little more satisfied than those in the unguided condition. Richards and Timulak [25] documented that most participants in both groups found the treatment helpful, even though there was a nonsignificant trend showing that participants in the unguided condition found the treatment easy to use and were more likely to report lasting effects than participants in the guided condition.

Although results give evidence of cCBT for depression being highly accepted, it should be noted that several studies do not give an exact definition of their object of investigation. As a consequence, the studies' stated objective is not in accordance with the measures that were used to examine the user acceptance of cCBT. These questionnaires did not refer explicitly to acceptability, satisfaction, or experiences but target related constructs such as ease of use [25,49], usefulness [41,42], or credibility or expectancy [26].

Discussion

Summary of Evidence

We intended to conduct a comprehensive review of studies regarding the user acceptance of cCBT for depression, updating the findings of Kaltenthaler et al [21] who aimed to assess studies regarding the acceptability to patients of cCBT for depression. Because user acceptance of a treatment may be a determinant for individuals to start and adhere to cCBT, the objective of this paper was to systematically evaluate studies that refer to the user acceptance of cCBT for depression in terms of acceptability, satisfaction, and experiences.

Corresponding with the findings of Kaltenthaler et al [21], the majority of the 29 reviewed studies reveal high or very high levels of user acceptance of cCBT programs. In addition to scientifically proven effectiveness of cCBT for depression [6-10], this result indicates a positive prognosis for future usage of cCBT programs.

When examining the user acceptance of cCBT for depression, most studies employed direct measures. Only a few studies made use of only indirect measures and consulted take-up, dropout, or completion rates for examining the user acceptance of cCBT [37,40,44]. In light of the fact that there was no information provided as to why participants do not start or continue a program, the validity of take-up, dropout, and completion rates is limited. Therefore, reasons that were not associated with the quality and appeal of a cCBT program, such as technical problems (eg, incompatible graphics software), personal reasons (eg, a lack of motivation because of a medical condition), or circumstances of research (eg, availability of an incentive) can lead to misinterpretations in terms of user acceptance. Kaltenthaler et al [21] had similar objections and concluded that the refusal to take part in a study regarding cCBT may show reluctance to enter a trial, rather than a dislike of cCBT. To examine the reasons why eligible persons who consented to participate in cCBT chose not to begin or drop out of the program, qualitative research efforts should be developed similar to those of Gerhards et al [39].

In general, the reported take-up rates for cCBT programs were wide ranging, making it difficult to draw comparisons with take-up rates for face-to-face CBT. However, the majority of studies reported dropout rates that are comparable with those reported for face-to-face CBT. In the RCT by Ekeblad et al [51], there was a dropout rate of 40%. Hans and Hiller [52] published a meta-analysis in which a dropout rate of 24.63% with a range from 0% to 68% was reported. Thimm and Antonsen [53] conducted a trial that revealed a dropout rate of 17.5%.

Upon closer examination of the reviewed studies, a number of methodological inaccuracies become apparent. Often no precise distinctions were made regarding the definition of acceptance, operationalization, and presentation of results. As a consequence, terms such as acceptance or acceptability, satisfaction, and usability were used interchangeably, although they can have different meanings [12,23,26,34,36,42,47]. Moreover, measures that did not correspond to the object of investigation were used [12,23,26,36,37,41,42]. These findings imply that research on acceptance is reflective of the fact that this is still a young field and, hence, there is a lack of precise definitions and adequate quantitative and qualitative measures, as these can only be realized over time.

These theoretical considerations are central to Kollmann [16] and Roger [17] who analyzed the construct of acceptance. Kollmann [16] defines acceptance as a combination of the inner reflection and the expectation formation (level of attitude), an adoption of the innovative product (level of action), and a voluntary problem-centered use of it (level of usage). This corresponds with the reflections of Rogers [17] who developed a model of stages in the innovation-decision process, which posits that individuals pass through an innovation-decision process, starting from first knowledge of an innovation such as cCBT, to forming an attitude toward the innovation, to deciding to adopt or reject, to implementing the new idea, and finally to confirming of this decision [17]. Thus, a person's decision about the engagement with an innovative computerized treatment is not an instantaneous act; rather, it is a process that occurs over time, comprising a series of actions and decisions [17]. The examination of all these levels or stages culminating in acceptance over time requires a number of well-matched measures that can be employed longitudinally.

The studies that made a comparison between guided and unguided cCBT programs regarding user acceptance revealed highly diverse results [24-26]. Further research is needed to shed more light on the user acceptance of guided and unguided cCBT programs.

Differentiated user perceptions of cCBT were central in the presentation of qualitative results by Gerhards et al [39] and Knowles et al [28]. Both gained deeper insights into the perspective of participants toward the user acceptance of cCBT for depression with the help of semistructured interviews. The findings were expressed in terms of motivators and barriers, as well as strengths and weaknesses. Therefore, these qualitative findings extend the knowledge gained in quantitative studies by providing cross-connections of participants' views and comprehensive insights in their experiences of these innovative treatments.

Strengths and Limitations

To our knowledge, this is the first review updating the state of the art regarding the user acceptance of cCBT for depression since Kaltenthaler et al [21] published their review about the patient acceptability of cCBT for depression in 2008. The major strength of this review is the comprehensive insight into the state of research regarding the user acceptance of cCBT by looking at various study types that give information about different approaches to ascertain the user acceptance of cCBT for depression.

There are a number of limitations to this review. The results of the studies provide a good overview of the user acceptance of cCBT for depression; however, they differ considerably in design, including sample characteristics, program features, and the condition under which treatment was offered. For example, four studies gave information on the user acceptance of cCBT reporting only on those participants who completed the treatment [29,32,33,50], whereas 18 studies analyzed data on user acceptance of cCBT also from noncompleters [12,24,26-28,30,31,34,35,37-40,43,44,47-49]. Seven studies did not provide any information as to whether completers or noncompleters or both had been included in analyses regarding the user acceptance of cCBT [23,25,36,41,42,45,46]. These different approaches to the consideration of participants for analysis make the studies difficult to compare.

There had been considerations to assess the quality of the studies formally. Since we included various study types ranging from RCTs to comparative trials to qualitative studies (see Multimedia Appendix 1), it was difficult to scrutinize them. Therefore, we decided to waive a quality assessment.

Furthermore, research on user acceptance is vulnerable to a selection bias because the process of accepting may already begin "before" using an innovative treatment, which means that people who have reservations regarding cCBT for depression may do not get involved in the first place. Moreover, it remains unknown if the refusal to participate in a study originates from reservations regarding cCBT or research itself. At the same time, research aspects may have an opposite unintended consequence; the program may encourage participation and adherence simply because it is being researched. Furthermore, user acceptance of cCBT for depression may be influenced by aspects associated with the user's medical condition. Thus, depressive mood, a loss of energy and drive as characteristics of depressive disorders may affect the motivation to start or adhere to cCBT. In addition to these aspects, the severity of symptoms and possible comorbidities are difficult to examine in terms of user acceptance.

Conclusions and Implications for Further Research

In conclusion, users of cCBT for depression experience the treatment as predominantly positive, which supports the potential benefit of innovative treatments such as cCBT. The preferred measures for examining the user acceptance in terms of acceptability, satisfaction, and experiences with cCBT were well-established questionnaires but principally study-specific developed questionnaires. Indirect measures such as completion, take-up, and dropout rates, as well as reasons for take-up and dropout were less common. However, there is considerable discrepancy regarding the objective's definition and operationalization.

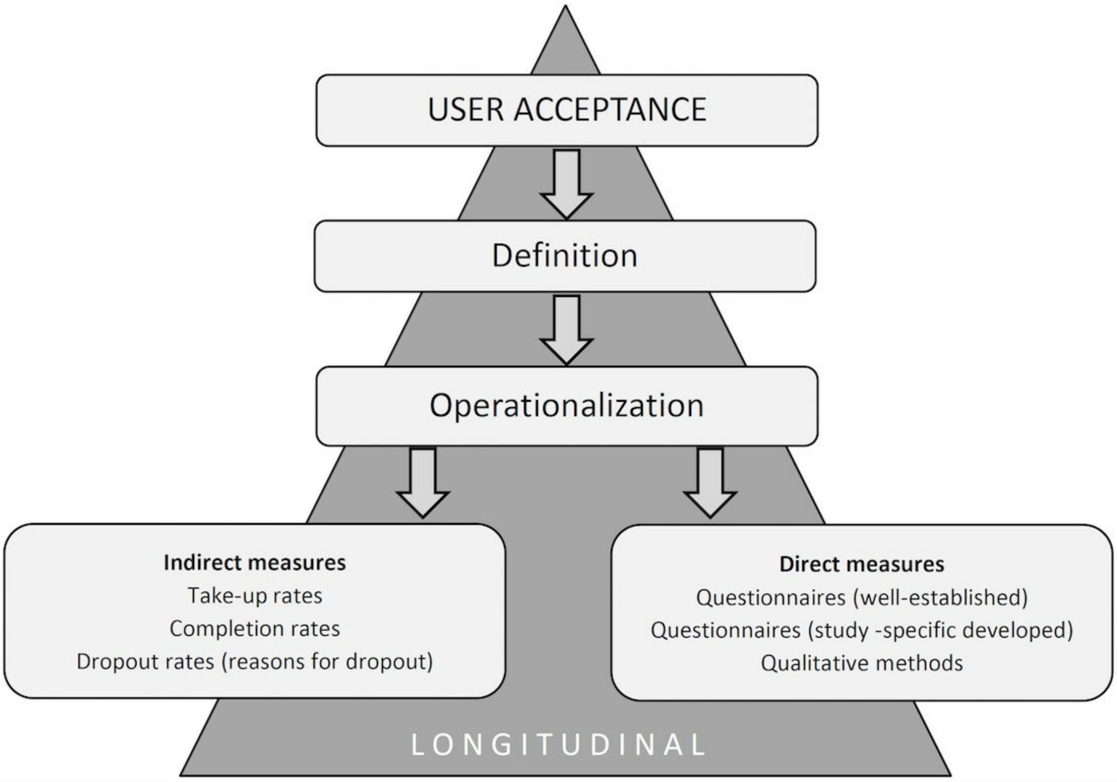
As can be seen in Figure 2, future research on user acceptance of cCBT should, therefore, include a theoretical framework and a definition of acceptance, adequate operationalization, and quantitative as well as qualitative data collection instruments that should be used in longitudinal and multidimensional approaches considering the stages of the process of acceptance.

The consideration of qualitative data is important since the accumulated material contains more details about the perspectives of trial participants than quantitative data does. Hence, in addition to take-up, completion, and dropout rates, it is important to learn about the reasons for take-up and dropout because one cannot be sure if discontinuing a treatment results from a negative attitude toward the treatment or other reasons such as those associated with research, technical, or personal circumstances.

A combination of quantitative and qualitative investigation examining expectations and experiences may prove beneficial. With the help of the juxtaposition of expectations and experiences, research on acceptance may fulfill its interpretation as a process in keeping with Kollmann's [16] notion of acceptance, including the levels of attitude, action, and usage as well as Rogers' [17] considerations regarding the five stages of the innovation-decision process. For this purpose, appropriate measures should be developed that are suitable for longitudinal studies. In general, the examination of user acceptance should be included alongside trials that focus basically on effectiveness.

In accordance with Kaltenthaler et al [21] and Waller and Gilbody [20], future research on user acceptance of innovative treatments such as cCBT should also include health care providers. Since physicians and therapists give therapy recommendations to their patients, it is important to learn more about their attitude toward cCBT.

Figure 2. Recommended examination of user acceptance.



Acknowledgments

This study was funded in affiliation with the German Research Foundation (DFG, grant number: LO 2171 / 1-1) and is published in affiliation with the German Federal Ministry of Education and Research (grant number: 01GY1613).

Conflicts of Interest

None declared.

Multimedia Appendix 1

Characteristics of the studies reviewed.

[PDF File (Adobe PDF File), 324KB - jmir_v19i9e309_app1.pdf]

Multimedia Appendix 2

Measures and results of the studies reviewed.

[PDF File (Adobe PDF File), 328KB - jmir_v19i9e309_app2.pdf]

Multimedia Appendix 3

Dropout and reasons for dropout of the studies reviewed.

[PDF File (Adobe PDF File), 283KB - jmir_v19i9e309_app3.pdf]

References

1. Whiteford HA, Ferrari AJ, Degenhardt L, Feigin V, Vos T. The global burden of mental, neurological and substance use disorders: an analysis from the global burden of disease Study 2010. PLoS One 2015;10(2):e0116820 [FREE Full text] [doi: 10.1371/journal.pone.0116820] [Medline: 25658103]
2. WHO. 2016. Depression URL: http://www.who.int/mediacentre/factsheets/fs369/en/ [accessed 2016-04-14] [WebCite Cache ID 6ntoaaxmB]
3. Hollon SD, Stewart MO, Strunk D. Enduring effects for cognitive behavior therapy in the treatment of depression and anxiety. Annu Rev Psychol 2006 Jan;57:285-315. [doi: 10.1146/annurev.psych.57.102904.190044] [Medline: 16318597]
4. Gega L, Marks I, Mataix-Cols D. Computer-aided CBT self-help for anxiety and depressive disorders: experience of a London clinic and future directions. J Clin Psychol 2004 Feb;60(2):147-157. [doi: 10.1002/jclp.10241] [Medline: 14724922]
5. Learmonth D, Trosh J, Rai S, Sewell J, Cavanagh K. The role of computer-aided psychotherapy within an NHS CBT specialist service. CPR 2008 Jun;8(2):117-123. [doi: 10.1080/14733140801976290]
6. 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]
7. Griffiths K, Farrer L, Christensen H. The efficacy of internet interventions for depression and anxiety disorders: a review of randomised controlled trials. Med J Aust 2010 Jun 07;192(11 Suppl):S4-11. [Medline: 20528707]
8. Andrews G, Cuijpers P, Craske MG, McEvoy P, Titov N. Computer therapy for the anxiety and depressive disorders is effective, acceptable and practical health care: a meta-analysis. PLoS One 2010;5(10):e13196 [FREE Full text] [doi: 10.1371/journal.pone.0013196] [Medline: 20967242]
9. Eells TD, Barrett MS, Wright JH, Thase M. Computer-assisted cognitive-behavior therapy for depression. Psychotherapy (Chic) 2014 Jun;51(2):191-197. [doi: 10.1037/a0032406] [Medline: 24059735]
10. Littlewood E, Duarte A, Hewitt C, Knowles S, Palmer S, Walker S, et al. A randomised controlled trial of computerised cognitive behaviour therapy for the treatment of depression in primary care: the Randomised Evaluation of the Effectiveness and Acceptability of Computerised Therapy (REEACT) trial. Health Technol Assess 2015 Dec;19(101):171 [FREE Full text] [doi: 10.3310/hta191010] [Medline: 26685904]
11. Sikorski C, Luppa M, Kersting A, König H, Riedel-Heller SG. [Computer-aided cognitive behavioral therapy for depression]. Psychiatr Prax 2011 Feb;38(2):61-68. [doi: 10.1055/s-0030-1248575] [Medline: 20972949]
12. 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]
13. Riedel-Heller SG. [Computer-based therapies: a supplement in modern psychotherapy? - pro]. Psychiatr Prax 2017 Jan;44(1):10-11. [doi: 10.1055/s-0042-115320] [Medline: 28049219]
14. Pressmar DB. Zur Akzeptanz von computergestützten Planungssystemen. In: Krallmann H, editor. Betriebs- und Wirtschaftsinformatik: Unternehmensplanung und -steuerung in den 80er Jahren. Berlin, Heidelberg: Springer; 1982:324-348.
15. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 1989 Sep;13(3):319. [doi: 10.2307/249008]
16. Kollmann T. Akzeptanz innovativer Nutzungsgüter: Konsequenzen für die Einführung von Telekommunikations- und Multimediasystemen. Wiesbaden: Gabler; 1998.
17. Rogers E. Diffusion of innovations. New York: Free Press; 1995.

18. Titov N. Status of computerized cognitive behavioural therapy for adults. Aust N Z J Psychiatry 2007 Feb;41(2):95-114. [doi: 10.1080/00048670601109873] [Medline: 17464688]
19. Vallury KD, Jones M, Oosterbroek C. Computerized cognitive behavior therapy for anxiety and depression in rural areas: a systematic review. J Med Internet Res 2015;17(6):e139 [FREE Full text] [doi: 10.2196/jmir.4145] [Medline: 26048193]
20. Waller R, Gilbody S. Barriers to the uptake of computerized cognitive behavioural therapy: a systematic review of the quantitative and qualitative evidence. Psychol Med 2009 May;39(5):705-712. [doi: 10.1017/S0033291708004224] [Medline: 18812006]
21. Kaltenthaler E, Sutcliffe P, Parry G, Beverley C, Rees A, Ferriter M. The acceptability to patients of computerized cognitive behaviour therapy for depression: a systematic review. Psychol Med 2008 Nov;38(11):1521-1530. [doi: 10.1017/S0033291707002607] [Medline: 18205964]
22. 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. [Medline: 19622511]
23. Danaher BG, Milgrom J, Seeley JR, Stuart S, Schembri C, Tyler MS, et al. MomMoodBooster web-based intervention for postpartum depression: feasibility trial results. J Med Internet Res 2013 Nov;15(11):e242 [FREE Full text] [doi: 10.2196/jmir.2876] [Medline: 24191345]
24. 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 Dec;40(4):251-266. [doi: 10.1080/16506073.2011.616531] [Medline: 22060248]
25. Richards D, Timulak L. Satisfaction with therapist-delivered vs. self-administered online cognitive behavioural treatments for depression symptoms in college students. Br J Guid Counc 2013 Apr;41(2):193-207. [doi: 10.1080/03069885.2012.726347]
26. Titov N, Andrews G, Davies M, McIntyre K, Robinson E, Solley K. Internet treatment for depression: a randomized controlled trial comparing clinician vs. technician assistance. PLoS One 2010;5(6):e10939 [FREE Full text] [doi: 10.1371/journal.pone.0010939] [Medline: 20544030]
27. Hind D, O'Cathain A, Cooper CL, Parry GD, Isaac CL, Rose A, et al. The acceptability of computerised cognitive behavioural therapy for the treatment of depression in people with chronic physical disease: a qualitative study of people with multiple sclerosis. Psychol Health 2010 Jul;25(6):699-712. [doi: 10.1080/08870440902842739] [Medline: 20204955]
28. Knowles SE, Lovell K, Bower P, Gilbody S, Littlewood E, Lester H. Patient experience of computerised therapy for depression in primary care. BMJ Open 2015 Nov 30;5(11):e008581. [doi: 10.1136/bmjopen-2015-008581] [Medline: 26621513]
29. Ahmedani BK, Crotty N, Abdulhak MM, Ondersma SJ. Pilot feasibility study of a brief, tailored mobile health intervention for depression among patients with chronic pain. Behav Med 2015;41(1):25-32. [doi: 10.1080/08964289.2013.867827] [Medline: 24313728]
30. Berman MI, Buckey JC, Hull JG, Linardatos E, Song SL, McLellan RK, et al. Feasibility study of an interactive multimedia electronic problem solving treatment program for depression: a preliminary uncontrolled trial. Behav Ther 2014 May;45(3):358-375 [FREE Full text] [doi: 10.1016/j.beth.2014.02.001] [Medline: 24680231]
31. Boeschoten RE, Nieuwenhuis MM, van Oppen P, Uitdehaag BM, Polman CH, Collette EH, et al. Feasibility and outcome of a web-based self-help intervention for depressive symptoms in patients with multiple sclerosis: A pilot study. J Neurol Sci 2012 Apr;315(1-2):104-109. [doi: 10.1016/j.jns.2011.11.016]
32. Burns MN, Begale M, Duffecy J, Gergle D, Karr CJ, Giangrande E, et al. Harnessing context sensing to develop a mobile intervention for depression. J Med Internet Res 2011;13(3):e55 [FREE Full text] [doi: 10.2196/jmir.1838] [Medline: 21840837]
33. Cartreine JA, Locke SE, Buckey JC, Sandoval L, Hegel MT. Electronic problem-solving treatment: description and pilot study of an interactive media treatment for depression. JMIR Res Protoc 2012 Sep 25;1(2):e11. [doi: 10.2196/resprot.1925]
34. Choi I, Zou J, Titov N, Dear BF, Li S, Johnston L, et al. Culturally attuned Internet treatment for depression amongst Chinese Australians: a randomised controlled trial. J Affect Disord 2012 Feb;136(3):459-468. [doi: 10.1016/j.jad.2011.11.003] [Medline: 22177742]
35. de Graaf LE, Huibers MJ, Riper H, Gerhards SA, Arntz A. Use and acceptability of unsupported online computerized cognitive behavioral therapy for depression and associations with clinical outcome. J Affect Disord 2009 Aug;116(3):227-231. [doi: 10.1016/j.jad.2008.12.009] [Medline: 19167094]
36. Dear BF, Zou J, Titov N, Lorian C, Johnston L, Spence J, et al. Internet-delivered cognitive behavioural therapy for depression: a feasibility open trial for older adults. Aust N Z J Psychiatry 2013 Feb;47(2):169-176. [doi: 10.1177/0004867412466154] [Medline: 23152358]
37. Dimidjian S, Beck A, Felder JN, Boggs JM, Gallop R, Segal ZV. Web-based mindfulness-based cognitive therapy for reducing residual depressive symptoms: an open trial and quasi-experimental comparison to propensity score matched controls. Behav Res Ther 2014 Dec;63:83-89. [doi: 10.1016/j.brat.2014.09.004] [Medline: 25461782]
38. Geraedts AS, Kleiboer AM, Wiezer NM, Cuijpers P, van Mechelen W, Anema JR. Feasibility of a worker-directed web-based intervention for employees with depressive symptoms. Internet Interv 2014 Jul;1(3):132-140. [doi: 10.1016/j.invent.2014.07.001]

39. Gerhards SA, Abma TA, Arntz A, de GL, Evers SM, Huibers MJ, et al. Improving adherence and effectiveness of computerised cognitive behavioural therapy without support for depression: a qualitative study on patient experiences. J Affect Disord 2011 Mar;129(1-3):117-125. [doi: 10.1016/j.jad.2010.09.012] [Medline: 20889214]
40. Kay-Lambkin F, Baker A, Lewin T, Carr V. Acceptability of a clinician-assisted computerized psychological intervention for comorbid mental health and substance use problems: treatment adherence data from a randomized controlled trial. J Med Internet Res 2011 Jan 27;13(1):e11. [doi: 10.2196/jmir.1522]
41. Kok G, Bockting C, Burger H, Smit F, Riper H. Mobile cognitive therapy: adherence and acceptability of an online intervention in remitted recurrently depressed patients. Internet Interv 2014 Apr;1(2):65-73. [doi: 10.1016/j.invent.2014.05.002]
42. Lintvedt OK, Griffiths KM, Sørensen K, Østvik AR, Wang CEA, Eisemann M, et al. Evaluating the effectiveness and efficacy of unguided internet-based self-help intervention for the prevention of depression: a randomized controlled trial. Clin Psychol Psychother 2013;20(1):10-27. [doi: 10.1002/cpp.770] [Medline: 21887811]
43. Lucassen MF, Merry SN, Hatcher S, Frampton CM. Rainbow SPARX: a novel approach to addressing depression in sexual minority youth. Cogn Behav Pract 2015 May;22(2):203-216. [doi: 10.1016/j.cbpra.2013.12.008]
44. McMurchie W, Macleod F, Power K, Laidlaw K, Prentice N. Computerised cognitive behavioural therapy for depression and anxiety with older people: a pilot study to examine patient acceptability and treatment outcome. Int J Geriatr Psychiatry 2013 Nov;28(11):1147-1156. [doi: 10.1002/gps.3935] [Medline: 23495124]
45. Merry SN, Stasiak K, Shepherd M, Frampton C, Fleming T, Lucassen MF. The effectiveness of SPARX, a computerised self help intervention for adolescents seeking help for depression: randomised controlled non-inferiority trial. BMJ 2012;344:e2598 [FREE Full text] [Medline: 22517917]
46. O'Mahen HA, Woodford J, McGinley J, Warren FC, Richards DA, Lynch TR, et al. Internet-based behavioral activation--treatment for postnatal depression (Netmums): a randomized controlled trial. J Affect Disord 2013 Sep 25;150(3):814-822. [doi: 10.1016/j.jad.2013.03.005] [Medline: 23602514]
47. Perini S, Titov N, Andrews G. Clinician-assisted Internet-based treatment is effective for depression: randomized controlled trial. Aust N Z J Psychiatry 2009 Jun;43(6):571-578. [doi: 10.1080/00048670902873722] [Medline: 19440890]
48. Schneider J, Sarrami FP, Grime P, Thornicroft G. Acceptability of online self-help to people with depression: users' views of MoodGYM versus informational websites. J Med Internet Res 2014 Mar;16(3):e90 [FREE Full text] [doi: 10.2196/jmir.2871] [Medline: 24681717]
49. Sheeber LB, Seeley JR, Feil EG, Davis B, Sorensen E, Kosty DB, et al. Development and pilot evaluation of an Internet-facilitated cognitive-behavioral intervention for maternal depression. J Consult Clin Psychol 2012 Oct;80(5):739-749 [FREE Full text] [doi: 10.1037/a0028820] [Medline: 22663903]
50. Stasiak K, Hatcher S, Frampton C, Merry SN. A pilot double blind randomized placebo controlled trial of a prototype computer-based cognitive behavioural therapy program for adolescents with symptoms of depression. Behav Cogn Psychother 2014 Jul;42(4):385-401. [doi: 10.1017/S1352465812001087] [Medline: 23253641]
51. Ekeblad A, Falkenström F, Andersson G, Vestberg R, Holmqvist R. Randomized trial of interpersonal psychotherapy and cognitive behavioral therapy for major depressive disorder in a community-based psychiatric outpatient clinic. Depress Anxiety 2016 Dec;33(12):1090-1098. [doi: 10.1002/da.22495] [Medline: 27029912]
52. Hans E, Hiller W. Effectiveness of and dropout from outpatient cognitive behavioral therapy for adult unipolar depression: a meta-analysis of nonrandomized effectiveness studies. J Consult Clin Psychol 2013 Feb;81(1):75-88. [doi: 10.1037/a0031080] [Medline: 23379264]
53. Thimm JC, Antonsen L. Effectiveness of cognitive behavioral group therapy for depression in routine practice. BMC Psychiatry 2014 Oct 21;14:292 [FREE Full text] [doi: 10.1186/s12888-014-0292-x] [Medline: 25330912]

Abbreviations

ARM: Agnes-Davies Relationship Measure
AST: Acceptability of Self-Guided Treatment Questionnaire
BDI: Beck depression inventory
BL: baseline
CBT: cognitive behavioral therapy
cCBT: computerized cognitive behavioral therapy
CEQ: Credibility or Expectancy Questionnaire
CSQ-8: Client Satisfaction Questionnaire
FU: follow-up
iBA: Internet-based behavioral activation
iCBT: Internet-based cognitive behavioral therapy
IG: intervention group
MS: multiple sclerosis
PICO: Patient, problem, or population; Intervention; Comparison, control, or comparator; Outcome

PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RCT: randomized controlled trial
SD: standard deviation
TAI: therapy attitude inventory
TAM: technology acceptance model
TAU: treatment-as-usual
VAS: visual analogue scale
ZUF-8: Fragebogen zur Patientenzufriedenheit (German version of CSQ-8)

Edited by G Eysenbach; submitted 10.03.17; peer-reviewed by M Ashford, J Apolinário-Hagen, TR Soron, S Langrial, N Bragazzi; comments to author 20.04.17; revised version received 09.06.17; accepted 07.07.17; published 13.09.17.

Please cite as:
Rost T, Stein J, Löbner M, Kersting A, Luck-Sikorski C, Riedel-Heller SG
User Acceptance of Computerized Cognitive Behavioral Therapy for Depression: Systematic Review
J Med Internet Res 2017;19(9):e309
URL: http://www.jmir.org/2017/9/e309/
doi:10.2196/jmir.7662
PMID:28903893

Original Paper

Personalized Hypertension Management Using Patient-Generated Health Data Integrated With Electronic Health Records (EMPOWER-H): Six-Month Pre-Post Study

Nan Lv[1], PhD; Lan Xiao[1], PhD; Martha L Simmons[2], MD, PhD; Lisa G Rosas[1,3], MPH, PhD; Albert Chan[2], MS, MD, FAAFP; Martin Entwistle[2,4], MBChB, MBA, FRCSEd

[1]Palo Alto Medical Foundation Research Institute, Palo Alto, CA, United States
[2]Office of Patient Experience, Sutter Health, Mountain View, CA, United States
[3]Department of Medicine, Stanford University School of Medicine, Palo Alto, CA, United States
[4]Ares Health Solutions, Pasadena, CA, United States

Corresponding Author:
Martin Entwistle, MBChB, MBA, FRCSEd
Ares Health Solutions
Suite 202
445 S. Los Robles Ave
Pasadena, CA, 91101
United States
Phone: 1 858 717 0465
Email: ment@areshs.com

Abstract

Background: EMPOWER-H (Engaging and Motivating Patients Online With Enhanced Resources-Hypertension) is a personalized-care model facilitating engagement in hypertension self-management utilizing an interactive Web-based disease management system integrated with the electronic health record. The model is designed to support timely patient-provider interaction by incorporating decision support technology to individualize care and provide personalized feedback for patients with chronic disease. Central to this process were patient-generated health data, including blood pressure (BP), weight, and lifestyle behaviors, which were uploaded using a smartphone.

Objective: The aim of this study was to evaluate the program among patients within primary care already under management for hypertension and with uncontrolled BP.

Methods: Using a 6-month pre-post design, outcome measures included office-measured and home-monitored BP, office-measured weight, intervention contacts, diet, physical activity, smoking, knowledge, and health-related quality of life.

Results: At 6 months, 55.9% of participants (N=149) achieved office BP goals (<140/90 mm Hg; *P*<.001) and 86.0% achieved clinically meaningful reduction in office BP (reduction in systolic BP [SBP] ≥5 mm Hg or diastolic BP [DBP] ≥3 mm Hg). At baseline, 25.2% of participants met home BP goals (<135/85 mm Hg), and this percentage significantly increased to 71.4% (*P*<.001) at 6 months. EMPOWER-H also significantly reduced both office and home SBP and DBP, decreased office-measured weight and consumption of high-salt and high-fat foods (all *P*<.005), and increased intake of fruit and vegetables, minutes of aerobic exercise, and hypertension knowledge (all *P*<.05). Patients with higher home BP upload frequencies had significantly higher odds of achieving home BP goals. Patients receiving more total intervention, behavioral, pharmaceutical contacts had significantly lower odds of achieving home BP goals but higher improvements in office BP (all *P*<.05).

Conclusions: EMPOWER-H significantly improved participants' office-measured and home-monitored BP, weight, and lifestyle behaviors, suggesting that technologically enabled BP home-monitoring, with structured use of patient-generated health data and a personalized care-plan facilitating patient engagement, can support effective clinical management. The experience gained in this study provides support for the feasibility and value of using carefully managed patient-generated health data in the day-to-day clinical management of patients with chronic conditions. A large-scale, real-world study to evaluate sustained effectiveness, cost-effectiveness, and scalability is warranted.

(J Med Internet Res 2017;19(9):e311) doi:10.2196/jmir.7831

KEYWORDS

electronic health records; disease management; hypertension; patient participation; blood pressure; patient-centered care; home blood pressure monitoring

Introduction

Hypertension is a major risk factor for cardiovascular disease (CVD), with an estimated annual cost of US $46 billion in the United States [1,2]. In national data, 29% of American adults have been diagnosed with hypertension and among those, only 52% have adequately controlled blood pressure (BP) under the recommended level of 140/90 mm Hg [3]. An additional 28% of Americans have prehypertension, a precursor of hypertension characterized by BP of 120/80 mm Hg to 139/89 mm Hg [4]. Furthermore, the problem is growing; it is estimated that 41% of American adults will have hypertension by 2030, with an estimated annual cost of US $274 billion [1]. Although common and costly, hypertension is preventable and modifiable through promotion of a healthy lifestyle, such as weight loss, healthy diet, and physical activity, or improving medication adherence.

Health policies have begun to shift away from episodic management of individual patients toward managing and paying for ongoing health care services that drive engagement and achieve valuable outcomes for groups of patients [5]. Population health management, the proactive application of strategies and interventions to defined groups of patients in an effort to improve health efficiently and at the lowest necessary cost [6], has been shown to be effective and cost-effective for conditions such as hypertension [7-9]. Population health management to prevent and control hypertension requires scalable and sustainable lifestyle interventions. Pragmatic technology-assisted approaches offered by the existing health care infrastructure may facilitate patient self-management and increase the potential for widespread reach and adoption, resulting in improved long-term effectiveness and a shift toward a population-based management model. Growing evidence [10,11] suggests that new approaches including technology-assisted clinical tools can both increase access and decrease cost for primary care–based hypertension prevention and management programs that traditionally place a heavy burden on personnel and resources. For example, technologies such as the Internet-based home BP telemonitoring can increase patient engagement and alert out-of-limit BP readings, which is a proven strategy for improved BP control [12-14]. Such technology-assisted programs have the potential to assist patients in self-managing their chronic conditions, including support for lifestyle changes, thus reducing reliance on health care personnel. Whereas evidence suggests that technology-assisted interventions in health care settings for hypertension prevention and control are potentially scalable and cost-effective, best practices remain unknown.

The Engaging and Motivating Patients Online With Enhanced Resources (EMPOWER) program is an innovative care delivery model for chronic disease management utilizing an interactive Web-based system integrated with the electronic health record (EHR). The model is designed to support timely patient-provider interaction by incorporating decision support technology to individualize care. The program also supports patient self-management and engagement by providing real-time feedback on progress against clinical goals to facilitate care-plan engagement. The model is designed to be generalizable to any chronic condition where behavior change has a direct impact on health outcomes. In this pilot study, we sought to evaluate the EMPOWER-Hypertension (EMPOWER-H) program among patients with uncontrolled BP in the context of primary care, using data from home BP monitoring as a major input to clinical decision making.

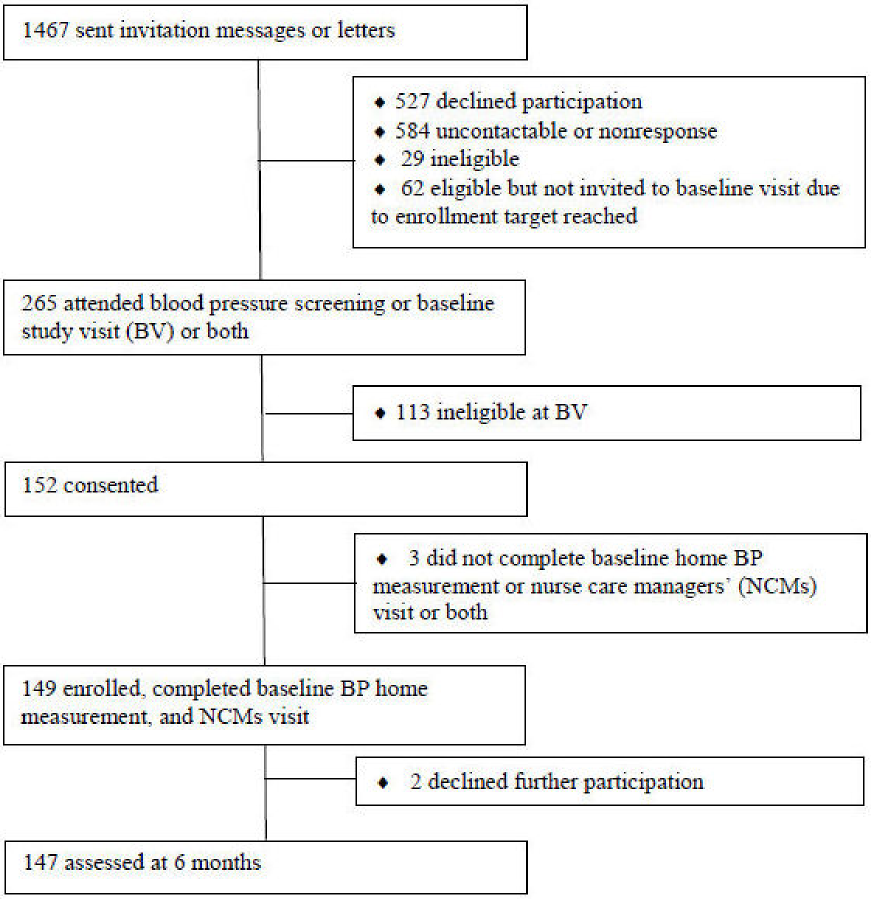
Methods

We evaluated the EMPOWER-H intervention using a pre-post design. An institutional review board of Palo Alto Medical Foundation (PAMF), an affiliate of Sutter Health, approved the study. All participants provided written informed consent.

Study Participants

Participants were recruited (March 2012 to June 2012) from two clinical sites at PAMF, a large ambulatory health care system. Patients were eligible to participate in the pilot study if they were in the age range of 35 to 75 years, had uncontrolled BP, and had been treated within PAMF during the previous 6 months. Participants were considered to have uncontrolled BP if (1) the average of 2 BP readings taken within 6 months of the study start date and separated by at least 2 weeks was 140 mm Hg to 175 mm Hg for SBP or 90 mm Hg to 110 mm Hg for DBP or both and (2) BP was within this same range when taken by the research staff during an in-person baseline visit. Exclusion criteria included having serious medical conditions (eg, diabetes, renal failure type III or IV, coronary artery disease, or stroke), requiring management of psychiatric issues, being under primary care management for a diagnosis of hypertension for less than 6 months, participating in other research projects, or having special life circumstances (eg, pregnancy, planned relocation, past or planned gastric bypass, or having no insurance). Of the 1467 patients contacted for recruitment, 527 patients declined participation, 142 were ineligible, 584 were not contactable or nonresponsive, 62 were interested but not enrolled as the study reached full enrollment before their initial study visit, and 3 consented but did not complete baseline home BP measurement or nurse care managers' (NCMs) visit (first in-person intervention visit). This process yielded the target enrollment of 149 eligible and consenting participants (Figure 1).

Figure 1. Consolidated Standards of Reporting Trials (CONSORT) chart.



Recruitment and Screening Process

Recruitment began by querying the EHR to identify patients who met the inclusion criteria. Upon identification, a clinical study coordinator conducted chart reviews to verify BP readings during the previous 6 months, as well as other exclusion criteria. A list of potentially eligible patients was provided to primary care physicians (PCPs) for final approval. After approval, PCPs contacted their eligible patients either via a secure Web-based patient portal (My Health Online) message, or by letter and encouraged them to participate in the study. The Web-based patient portal is an integral part of the PAMF EHR. Patients can log in to message their medical providers, view lab results, and receive information about appointments. The initial communication included instructions on accessing the study enrollment site where the patient could read a description of the study, provide consent online, complete a self-screening questionnaire (eg, age, and pregnancy) and, if eligible, complete a baseline questionnaire.

For patients who were eligible per the initial Web-based screening, trained research staff arranged an in-person baseline visit where written consent was attained, and BP was measured in the office using the patient's home BP monitor according to standardized protocols [15] to confirm that the patient met the BP inclusion criteria.

Intervention

EMPOWER-H was a 6-month intervention delivered by a care team, including 2 NCMs, a registered dietitian (RD), and a pharmacist for consultation. The intervention was based on theoretical constructs such as perceived severity of a health threat; relevant values; self-efficacy; perceived barriers to action, as described in the theory of planned behavior [16]; health belief model [17]; and social cognitive theory [18]. It used an enhanced version of the Web-based disease management system as deployed to support EMPOWER-Diabetes (EMPOWER-D) [19], a precursor of EMPOWER-H. EMPOWER-H included (1) a wireless BP monitor that transmitted home BP readings to PAMF's EHR and the EMPOWER system; (2) a smartphone (Apple iPhone 3) with 2 study apps (described below); (3) a comprehensive dashboard of the status of a patient's personalized action plan, treatment goals, and self-monitoring data, available directly from within PAMF's Web-based patient portal (My Health Online); (4) a pedometer for monitoring steps; (5) Web-based messaging system for communicating between patients and members of the care team; (6) NCMs assisted by RD for nutrition and weight management and pharmacist to provide consultation and make medication changes; and (7) patient-specific text and video educational nuggets (Multimedia Appendix 1). On the smartphone, the Numera app allowed wireless transmission of the home BP data to the smartphone and then to the EHR and EMPOWER system. The EMPOWER-H app displayed the patient-generated home data, allowed visualization and tracking of personal goals, and provided access to educational nuggets.

At the baseline visit, research staff provided eligible patients with their study tools as described above, including the BP monitor with appropriate sized cuff, a pedometer, and a smartphone, along with instructions for use. Research staff also introduced patients to their personal dashboard and provided educational handouts and actions to take in case of very high or very low BP measurements. Patients were instructed to measure and upload BP readings in the morning and evening at regular times or events (eg, before breakfast or dinner) for at least 3 days a week over the following 7 days (baseline home BP monitoring) and to continue the same pattern of uploading during the entire 6-month intervention. They were also encouraged to use their pedometer and upload daily step count. A scale that allowed wireless upload of weight was provided to selected patients at a later visit, based on clinical indications determined by an RD.

Participants had a first in-person visit with an NCM as soon as possible after the 7-day baseline home BP collection period. This visit included the following activities: (1) addressing participants' questions; (2) providing education about cardiovascular risks; (3) reviewing how to use study tools; (4) setting 2 to 3 small attainable goals utilizing motivational interviewing techniques; (5) visiting the Web-based dashboard, entering the personalized goals, and showing patients how to enter and view their data; and (6) developing a personalized BP management plan (eg, frequency of contacts for check-ins, goal-setting, and data upload) informed by the 7-day baseline home BP. Home BP goals were set at American Heart Association's (AHA) recommended levels of <135 mm Hg for SBP and <85 mm Hg for DBP [20]. The NCM encouraged all patients to have a separate visit with an RD to discuss diet history, current dietary habits, and recommended changes. The vast majority (95.3%) of patients had at least one contact with an RD.

The NCM contacted each patient via EHR secure messaging or telephone or both as soon as possible 1 week after the first visit to assess engagement, challenges with both the technology and clinical management, as well as to check BP control and modify the BP management plan if needed.

During the 6-month intervention, patients were monitored daily using a provider dashboard which used both a 14-day average of home BP data and frequency of BP uploading (as a measure of engagement) to prioritize patients into high-risk (red), medium-risk (yellow), or low-risk (green) category (Multimedia Appendix 2). The dashboard also provided alerts if individual home BP measurements were at a critical level (SBP ≤90 or ≥160; DBP ≤55 or ≥100). NCMs who worked with an RD and pharmacist strictly followed the hypertension clinical management protocol that was developed for registered nurse (RN) management of hypertension based on the 7th report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure [2] and approved by PAMF nursing governance to provide continuous patient support. The protocol provides narrowly defined parameters within which an NCM may make adjustments to medical and therapeutic management, including limits on how frequently management adjustments may be made. Situations that fall outside the defined areas for modifications to management must be referred to the treating physician, a PCP, or a specialist for determination of action. When further assistance with medication adjustment was needed, the pharmacist reviewed charts and consulted with an NCM.

An NCM or RD had a library of Web-based education handouts and feedback message nuggets (texts, links to Web pages, or short videos) that could be sent via secure messaging on the Web-based patient portal throughout the intervention. NCMs and RD primarily used secure messaging to provide regular pertinent electronic messages and timely feedback about participants' clinical variables (eg, home BP readings, medication doses, weight, and steps). In addition to these Web-based learning opportunities, participants were invited to participate in other engagement activities throughout the intervention, including healthy shopping tours and cooking classes at a grocery store, educational and interactive webinars on behavioral change, a pedometer challenge, and a healthy recipe challenge.

Outcome Measures

The 149 enrolled participants were assessed at baseline and 147 (98.7%) were assessed at 6 months (Figure 1). All outcome assessors were trained to perform the measurements per standardized protocols and procedures (Multimedia Appendix 3).

We assessed the effectiveness of the intervention using BP measurements taken in the office by trained research staff (office BP) and taken at home by the patient (home BP) as part of self-monitoring for the intervention. Office BP was measured by trained research staff according to standardized protocols [15]. The primary outcome for this study was the percentage of participants achieving office BP goals (<140 mm Hg for SBP and <90 mm Hg for DBP) at 6 months as measured in controlled circumstances. Secondary BP outcomes included change in office BP, percentage of participants achieving clinically meaningful reductions in office BP (a drop in SBP of ≥5 mm Hg or a drop in DBP of ≥3 mm Hg) [21-24], change in home BP measurements, and percentage of participants achieving home BP goals (<135 mm Hg for SBP and <85 mm Hg for DBP) at 6 months. Baseline home BP was calculated as an average of BP self-monitored during the 7 days after baseline, and 6-month home BP was an average of BP self-monitored during the 7 days before the 6-month visit.

Other secondary outcomes included number of participants meeting weekly home BP monitoring frequency target (upload twice a day and 3 days in a week), body mass index (BMI), weight, dietary intake, physical activity, smoking status, hypertension knowledge, as well as health-related quality of life (HRQoL). Home-monitored BP data were extracted from the EHR that pulled data from the Numera database. Indicators of dietary intake included frequency of consuming fruit and vegetables, high-salt food, and high-fat food measured by scales adapted from the validated Block food screeners [25,26]. Physical activity was measured by the Stanford Exercise Behavior Scale [27]. Hypertension knowledge was measured using a 13-item knowledge questionnaire. HRQoL was measured by the Veterans RAND 12-Item Health Survey (VR-12) [28-31]. The VR-12 was developed from the Veterans RAND 36-Item Health Survey that was developed and modified from the original RAND version of the 36-item Health Survey version 1.0 (also known as the “MOS SF-36”).

Indicators of engagement in the intervention included total number of times patients uploaded BP measurement, pedometer step data, weight, stress (ie, stress level on a 5-point scale), and medication (ie, whether medication was taken as prescribed) to the Numera and EMPOWER-H apps. Other indicators of intervention engagement included participation in 6 additional activities designed to support education and motivation (ie, pedometer and recipe challenges, 2 cooking classes, and 2 learning webinars).

In addition, we tracked the number of intervention contacts (ie, total or patient-initiated intervention visits, My Health Online messages, phone messages, or phone calls on behavioral, pharmaceutical, laboratory, or technical issues).

Participant activation was measured by the 13-item Patient Activation Measure (PAM) assessing patient’s knowledge, skill, and confidence for self-management [32,33]. PAM score has a range of 0 to 100, with higher score indicating higher patient activation.

Statistical Analyses

To examine our primary outcome, we used a one sample proportion test to compare the percentage of participants achieving office BP goals (<140 mm Hg for SBP and <90 mm Hg for DBP) at 6 months with the 0% who were meeting goals at baseline. As supplementary analyses, we compared the percentage of participants achieving office BP goals at 6 months with the percentage of patients who achieved normal BP with usual care alone in the same clinics during a similar period (30%). Frequencies and percentages were calculated for categorical variables, and means and standard deviations (SD) were calculated for continuous variables. Study outcomes at baseline and 6 months were compared using paired *t* tests for continuous outcomes and McNemar tests on paired proportions for categorical outcomes. McNemar exact tests were performed for the outcomes with small counts within certain categories (eg, 20% of cells with expected frequencies <5). We used mixed model growth curve analysis for all available home-monitored BP during the study period and plotted the results according to time. We analyzed SBP and DBP as an outcome in two separate models; each included random intercept and random and fixed effects of slope and quadratic term of days from start. The quadratic term of days was removed in the final model for both home-monitored SBP and DBP because it was not significant. We used logistic regression and regression models to examine the bivariate associations of indicators of intervention engagement with the categorical and continuous BP outcomes, respectively.

All analyses were conducted in Statistical Analysis System (SAS) version 9.3 (SAS Institute Inc). Statistical significance was set at $P<.05$ (two-tailed). We powered this study on the primary outcome of percentage of participants achieving office BP goals (<140 mm Hg for SBP and <90 mm Hg for DBP) at 6 months. A sample of 149 had 80% power to detect at least 40% of participants achieving office BP goals at 6 months at alpha=.05 (two-sided), assuming at least 90% retention at 6 months based on our prior experiences. We based our estimates on the 30% transition rate from abnormal BP to normal BP without intervention using patient data from our EHR assessed from 2009 to 2010.

Results

Baseline Characteristics

Participants had a mean (SD) age of 62.2 (9.5) years, 51% were female, 73% were married or lived with partner, and 76% were non-Hispanic white (Table 1). Most participants completed at least some college, had an annual family income of >US $75,000, were full-time employees or retired, and were never or former smokers. At baseline, their mean (SD) for BMI was 28.7 (6.2) kg/m^2, SBP was 149.8 (9.8) mm Hg, DBP was 91.0 (8.0) mm Hg, and PAM score was 43.1 (6.1).

Table 1. Baseline characteristics of the study participants.

Characteristic	All participants (N=149)
Age in years, mean (SD[a])	62.2 (9.5)
Female, n (%)	76 (51.0)
Marital status, n (%), n=147	
Married or partner or domestic partnership	107 (72.8)
Divorced or separated or widowed	25 (17.0)
Single	15 (10.2)
Race or ethnicity, n (%), n=146	
Non-Hispanic white	111 (76.0)
Non-Hispanic black	9 (6.2)
Hispanic	4 (2.7)
Asian	18 (12.3)
Other	4 (2.7)
Education, n (%), n=147	
High school or GED[b] or less	5 (3.4)
College, 1-3 years	28 (19.1)
College, 4 years or more	44 (29.9)
Postgraduate	70 (47.6)
Annual family income in US dollars, n (%), n=111	
$35,000-<$75,000	19 (17.1)
$75,000-<$100,000	28 (25.2)
$100,000-<$125,000	18 (16.2)
$125,000-<$150,000	10 (9.1)
$150,000+	36 (32.4)
Employment status, n (%), n=147	
Full-time	67 (45.6)
Part-time	10 (6.8)
Homemaker	8 (5.5)
Retired	49 (33.3)
Unemployed or disabled or something else	13 (8.8)
Smoking status, n (%)	
Never smoked	101 (67.8)
Current smoker	8 (5.4)
Former smoker	40 (26.9)
Body mass index, kg/m^2, mean (SD), n=140	28.7 (6.2)
Systolic blood pressure, mm Hg, mean (SD), n=145	149.8 (9.8)
Diastolic blood pressure, mm Hg, mean (SD), n=145	91.0 (8.0)
Patient activation measure score[c], mean (SD)	43.1 (6.1)

[a] SD: standard deviation.

[b] GED: general education diploma.

[c] 13-item Patient Activation Measure (PAM) assessing patient’s knowledge, skill, and confidence for self-management has a range of 0 to 100. Higher score indicates higher patient activation.

Table 2. Changes in office-measured blood pressure (BP), home-monitored BP, and body mass index at 6 months.

Outcome variable	n[a]	Baseline	6 months	Change	*P* value
Office-measured BP [b]	149, 143, 143				
Participants achieving BP goals, %[e]		0	55.9	55.9	<.001
Systolic BP, mm Hg, mean (SD[d])		149.8 (9.8)	134.4 (14.0)	−15.2 (15.3)	<.001
Diastolic BP, mm Hg, mean (SD)		91.0 (8.0)	84.5 (8.6)	−6.4 (7.5)	<.001
Home-monitored BP	147, 133, 132				
Participants achieving BP goals, %[c]		25.2	71.4	46.2	<.001
Systolic BP, mm Hg, mean (SD)		138.4 (10.6)	126.7 (9.8)	−11.7 (11.5)	<.001
Diastolic BP, mm Hg, mean (SD)		84.5 (7.8)	78.6 (6.9)	−6.1 (6.5)	<.001
Body mass index, kg/m^2, mean (SD)	140, 136, 132	28.7 (6.2)	28.3 (6.0)	−0.4 (1.5)	.01
Weight, kg, mean (SD)	142, 138, 134	84.0 (21.7)	82.0 (20.7)	1.2 (4.5)	.002

[a] n is provided for baseline, 6 months, and change in order.

[b] BP: blood pressure.

[c] Office BP goals are <140 mm Hg for systolic and <90 mm Hg for diastolic.

[d] SD: standard deviation.

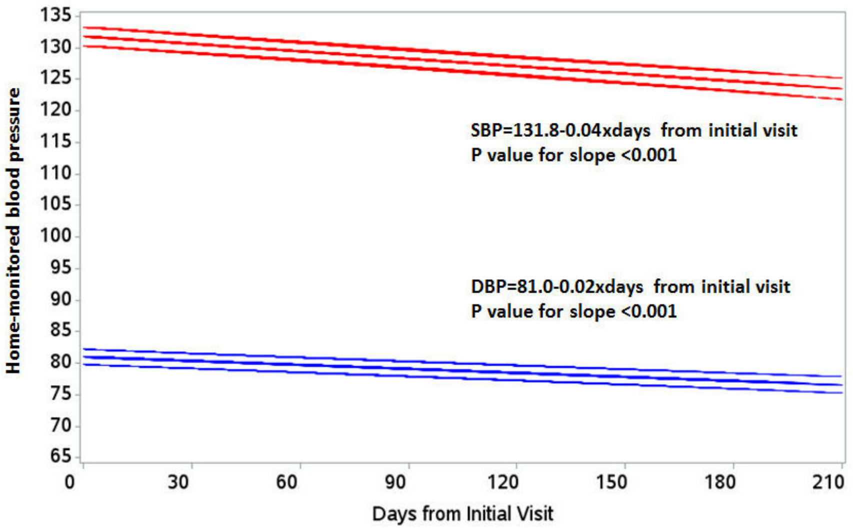
[e] Home BP goals are <135 mm Hg for systolic and <85 mm Hg for diastolic.

Changes in Clinical Outcomes

Table 2 shows that 55.9% of the participants achieved office BP goals (<140 mm Hg for SBP and <90 mm Hg for DBP) at 6 months (*P*<.001). Additionally, this was also significantly higher than 30%, which was achieved with usual care alone based on 2009-2010 patient data in EHR (*P*<.001). Paired *t* test results show that mean (SD) systolic office BP was significantly reduced from 149.8 (9.8) to 134.4 (14.0) mm Hg and mean (SD) diastolic office BP from 91.0 (8.0) to 84.5 (8.6) mm Hg (both *P*<.001). At 6 months, 86% of participants achieved clinically meaningful reduction in office BP (reduction in SBP ≥5 mm Hg or reduction in SBP ≥3 mm Hg).

Home-monitored BP measurements show that percent of participants achieving home BP goals (<135 mm Hg for SBP and <85 mm Hg for DBP) significantly increased from 25.2% to 71.4% (*P*<.001), with a mean (SD) reduction of 11.7 (11.5) mm Hg for systolic and 6.1 (6.5) mm Hg for diastolic home BP (both *P*<.001; Table 2). Figure 2 shows significant linear downward trends in both home-monitored SBP and DBP (both *P*<.001). Of 132 participants with both baseline and 6-month home BP measurements, 27 (20%) started and stayed within the home BP goal at 6 months, 67 (51%) were not at goal at baseline and were meeting goal at 6 months, 6 (5%) started within goal but were above goal at 6 months, and 32 (24%) started and stayed outside goal at 6 months (data not shown). Comparison between office and home BP shows that change in office and home BP at 6 months was significantly correlated (*P*<.001 for SBP and *P*=.005 for DBP; data not shown).

Figure 2. Predicted means and 95% CIs resulting from the random coefficient regression model of home-monitored blood pressures (N=149).



Compared with baseline, participants' BMI and body weight were significantly reduced (*P*=.01 for BMI and *P*=.002 for body weight; Table 2). Of the 149 participants, 121 (81.2%) received a wireless weight scale. There was no significant correlation between number of weight uploads and weight loss (*P*=.49)

Changes in Behavioral Outcomes

Participants significantly increased consumption frequency of fruit and vegetables (*P*=.01) and reduced consumption frequency of high-salt and high-fat foods (both *P*<.001; Table 3). Minutes of aerobic exercise per week significantly increased (*P*=.03), whereas minutes of stretching or strengthening remained the same (*P*=.91). Participants' hypertension knowledge also significantly improved (*P*<.001); however, smoking status and HRQoL measured by VR-12 remained unchanged (*P*>.05 for all; Table 3).

Correlation of Intervention Engagement With Changes in Office-Measured and Home-Monitored Blood Pressure (BP)

Almost half (44.3%) of participants did at least one of the 6 additional activities designed to support sustained program engagement (pedometer and recipe challenges, 2 cooking classes, and 2 learning webinars). None of the intervention engagement measures (ie, number of home BP, pedometer data, weight, stress, and medication uploads, challenges, optional events, and intervention contacts) were associated with achieving office BP goals or clinically meaningful office BP improvement at 6 months (Table 4). However, a higher number of total interventions, behavioral, pharmaceutical, and total patient-initiated intervention contacts were significantly associated with higher improvements in values of systolic or diastolic office BP or both.

Table 3. Changes in behavioral outcomes, knowledge, and quality of life at 6 months.

Outcome variable	n[a]	Baseline	6 months	Change	*P* value
Diet, times/week, mean (SD [b])	147, 140, 138				
Fruit and vegetables[c]		25.5 (10.5)	27.9 (11.3)	2.3 (10.8)	.01
High-salt food[d]		8.4 (5.1)	6.5 (3.8)	−1.8 (4.5)	<.001
High-fat food[e]		15.9 (8.9)	12.5 (7.0)	−3.3 (6.3)	<.001
Physical activity, minutes/week, mean (SD)	147, 140, 138				
Stretching or strengthening		51.7 (60.3)	53.0 (58.9)	−0.5 (55.7)	.91
Aerobic exercise		178.6 (132.4)	206.4 (126.2)	25.7 (133.1)	.03
Smoking status, %	149, 139, 139				.15
Never smoker		67.8	66.9		
Current smoker		5.4	2.9	N/A	
Former smoker		26.9	30.2		
Hypertension knowledge, mean (SD)[f]	147,140,138	11.3 (1.6)	12.3 (1.0)	1.0 (1.4)	<.001
VR-12, physical component score, mean (SD)[g]	147,140,138	49.7 (8.0)	50.2 (7.8)	0.5 (7.1)	.39
VR-12, mental component score, mean (SD)[g]	147,140,138	53.1 (7.2)	53.5 (8.3)	0.1 (8.6)	.90

[a] n is provided for baseline, 6 months, and change in order.

[b] SD: standard deviation.

[c] Sum of eating frequency (0.5=less than 1/week or never, 1=once a week, 2.5=2 to 3 times a week, 5=4 to 6 times a week, 7=once a day, and 14=2 or more times a day) for 7 fruit and vegetable subgroups (fruit juice; fresh, canned, or frozen fruit; vegetable juice; green salad; potatoes; vegetable soup or stew with vegetables; and any other vegetables, including string beans, peas, corn, broccoli, or any other kind).

[d] Sum of eating frequency (0.5=less than 1/week or never, 1=once a week, 2.5=2 to 3 times a week, 5=4 to 6 times a week, 7=once a day, and 14=2 or more times a day) for 6 high-salt food consuming and using behaviors (restaurant food; packaged snack foods such as chips, pretzels, popcorn, salted nuts; canned soups, canned vegetables, or frozen meals or TV dinners; cured or salted meats; add salt to your food at the table; add any of the following to your food when preparing meals or eating out: salt, mustard, pickles, relish, soy sauce, ketchup, meat tenderizer, or MSG).

[e] Sum of eating frequency (0.125=less than 1 time per month or never, 0.375=2 to 3 times a month, 1.5=1 to 2 times a week, 3.5=3 to 4 times a week, and 6=5+ times a week) for 17 high-fat foods or food subgroups (hamburgers, ground beef, meat burritos, tacos, enchiladas; pork chops, beef; fried chicken; hot dogs or Polish or Italian sausage, organ meats; cold cuts, lunch meats, ham; bacon or breakfast sausage; salad dressing; margarine, butter, lard or mayo spread on bread or potatoes; margarine, butter, lard or oil in cooking; eggs; pizza; cheese, cheese spread; whole milk or chocolate milk; French fries or fried potatoes; corn chips, potato chips, popcorn, crackers, peanuts; doughnuts, pastries, cake, cookies, pan dulce; ice cream).

[f] Hypertension knowledge was measured using a 13-item knowledge questionnaire. Higher score indicates better knowledge.

[h] Physical and mental component scores are calculated to report an overall measure of physical and mental functioning that is comparable among the surveys. These summary scales have been normalized in the US population (value=50). The higher score indicates better self-reported health-related quality of life.

Table 4. Bivariate associations of office-measured blood pressure (BP) with number of home BP, pedometer, weight, stress, and medication uploads and number of intervention contacts (n=143).

Intervention engagement measures	Achieving BP goals[a] OR[b] (95% CI)	Achieving clinically meaningful BP[c] improvement[d] OR (95% CI)	SBP[e] change[f] coefficient (95% CI)	DBP[g] change[f] coefficient (95% CI)
Number of weeks meeting home BP monitoring frequency target	1.04 (1.00-1.08)	0.98 (0.93-1.04)	−0.16 (−0.47 to 0.16)	0.16 (−0.09 to 0.22)
Number of other uploads				
Pedometer	0.99 (0.99-1.00)	0.99 (0.98-1.00)	0.11 (−0.05 to 0.07)	0.03 (−0.01 to 0.07)
Weight	1.00 (0.99-1.01)	0.99 (0.98-1.01)	0.03 (−0.05 to 0.11)	0.02 (−0.01 to 0.05)
Stress	1.00 (0.99-1.01)	0.998 (0.99-1.00)	0.001 (−0.04 to 0.04)	−0.005 (−0.02 to 0.01)
Medication	1.00 (1.00-1.01)	1.00 (0.99-1.01)	−0.01 (−0.05 to 0.04)	−0.01 (−0.03 to 0.01)
Number of challenges				
Pedometer	0.88 (0.45-1.72)	1.38 (0.51-3.77)	−2.39 (−7.49 to 2.72)	−0.68 (−3.20 to 1.85)
Recipe	1.33 (0.31-5.81)	1.13 (0.13-9.68)	−1.89 (−12.93 to 9.15)	1.80 (−3.64 to 7.24)
Optional events				
Web-based learning	1.26 (0.34-4.69)	0.54 (0.13-2.35)	3.64 (−6.17 to 13.44)	3.20 (−1.61 to 8.02)
Cooking classes	1.01 (0.40-2.55)	0.73 (0.22-2.35)	−0.25 (−7.36 to 6.86)	2.21 (−1.28 to 5.70)
Total number of intervention contacts	1.03 (0.99-1.07)	1.06 (0.99-1.13)	−0.42 (−0.71 to −0.14)[h]	−0.18 (−0.32 to −0.04)[h]
Behavioral contacts	1.07 (0.98-1.17)	1.05 (0.92-1.20)	−0.73 (−1.37 to −0.10)[i]	−0.14 (−0.45 to 0.18)
Pharmaceutical contacts	1.01 (0.90-1.12)	1.06 (0.89-1.26)	−0.68 (−1.49 to 0.13)	−0.40 (−0.80 to −0.001)[i]
Laboratory contacts	1.22 (0.91-1.64)	1.21 (0.76-1.92)	−0.71 (−2.8 to 1.41)	−0.53 (−1.57 to 0.51)
Technical contacts	1.09 (0.91-1.31)	1.28 (0.91-1.82)	−0.50 (−1.84 to 0.84)	−0.25 (−0.91 to 0.41)
Total number of patient-initiated intervention contacts	1.07 (0.98-1.18)	1.16 (0.98-1.37)	−0.98 (−1.62 to −0.34)[h]	−0.34 (−0.66 to −0.02)[i]

[a] Office blood pressure goals are <140 mm Hg for systolic and <90 mm Hg for diastolic.

[b] OR: odds ratio.

[c] BP: blood pressure.

[d] Reduction in systolic BP ≥5 mm Hg or reduction in diastolic BP ≥3 mm Hg.

[e] SBP: systolic blood pressure.

[f] 6-month office BP—baseline office BP.

[g] DBP: diastolic blood pressure.

[h] *P*<.01.

[i] *P*<.05.

The number of participants meeting home BP monitoring frequency target (uploading twice a day and 3 days in a week) gradually decreased during the course of the intervention (Figure 3). Table 5 shows that more weeks of meeting home BP upload frequency target was significantly associated with higher odds of achieving home BP goals (*P*=.02); however, a higher number of total intervention, behavioral, pharmaceutical contacts were significantly associated with lower odds of achieving home BP goals (all *P*<.001). None of the intervention engagement measures was associated with improvements in values of home BP.

Additionally, none of the office and home BP outcomes were associated with baseline and change in participant activation, as measured by the PAM score (data not shown).

Figure 3. The number of participants meeting weekly home blood pressure monitoring frequency target (upload twice a day and 3 days in a week).

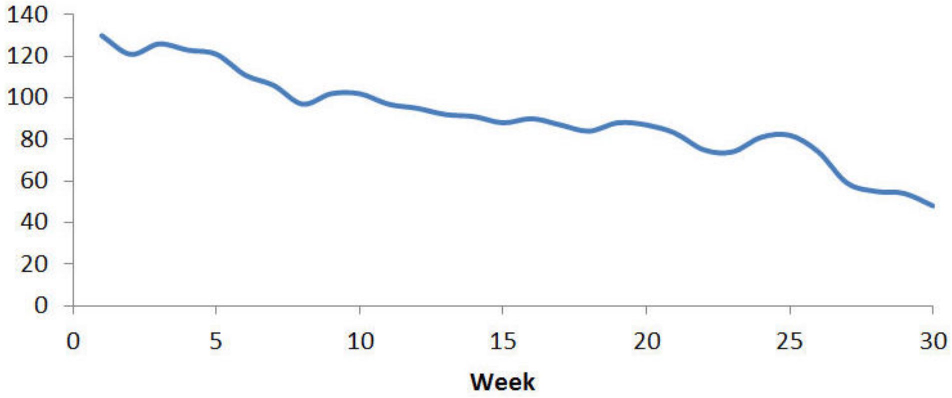


Table 5. Bivariate associations of home-monitored blood pressure (BP) with number of home BP, pedometer, weight, stress, and medication uploads and number of intervention contacts.

Intervention engagement measures	Achieving BP[a,b] goals[c] (n=133), OR[d] (95% CI)	SBP[e] change[f] (n=132), coefficient (95% CI)	DBP[g] change[f] (n=132), coefficient (95% CI)
Number of weeks meeting home BP monitoring frequency target	1.05 (1.01-1.10)[h]	−0.09 (−0.33 to 0.15)	−0.07 (−0.21 to 0.07)
Number of other uploads			
Pedometer	1.00 (1.00-1.01)	−0.02 (−0.07 to 0.03)	−0.01 (−0.03 to 0.02)
Weight	1.00 (0.99-1.01)	−0.06 (−0.121 to 0.00)	−0.003 (−0.04 to 0.03)
Stress	1.01 (1.00-1.02)	−0.01 (−0.04 to 0.02)	−0.01 (−0.02 to 0.01)
Medication	1.01 (1.00-1.01)	−0.01 (−0.04 to 0.02)	−0.004 (−0.02 to 0.01)
Number of challenges			
Pedometer	1.07 (0.50-2.30)	−0.80 (−4.81 to 3.21)	−0.38 (−2.66 to 1.90)
Recipe	1.21 (0.23-6.30)	−6.19 (−14.43 to 2.05)	−0.62 (−5.34 to 4.10)
Optional events			
Web-based learning	0.29 (0.07-1.17)	6.19 (−1.11 to 13.50)	3.78 (−0.36 to 7.93)
Cooking classes	0.52 (0.21-1.30)	2.21 (−1.28 to 5.70)	1.03 (−1.89 to 3.95)
Total number of intervention contacts	0.90 (0.85-0.94)[i]	−0.08 (−0.31 to 0.15)	−0.07 (−0.20 to 0.06)
Behavioral contacts	0.85 (0.76-0.94)[i]	−0.12 (−0.63 to 0.39)	−0.12 (−0.41 to 0.16)
Pharmaceutical contacts	0.77 (0.67-0.89)[i]	−0.12 (−0.77 to 0.52)	−0.27 (−0.63 to 0.09)
Laboratory contacts	0.81 (0.61-1.09)	−0.71 (−2.32 to 0.89)	−0.37 (−1.28 to 0.54)
Technical contacts	1.06 (0.85-1.32)	−0.08 (−1.16 to 1.00)	0.06 (−0.55 to 0.68)
Total number of patient-initiated intervention contacts	0.95 (0.87-1.04)	−0.42 (−0.93 to 0.10)	−0.10 (−0.39 to 0.20)

[a] BP: blood pressure.

[b] Baseline home BP was an average of BP self-monitored during the 7 days after baseline, and 6-month home BP was an average of BP self-monitored during the 7 days before the 6-month visit.

[c] Home BP goals are <135 mm Hg for systolic and <85 mm Hg for diastolic.

[d] OR: odds ratio.

[e] SBP: systolic blood pressure.

[f] 6-month home BP—baseline home BP.

[g] DBP: diastolic blood pressure.

[h] *P*<.05.

[i] *P*<.001.

Discussion

Principal Findings

This pre-post pilot study evaluated the EMPOWER-H program, a new personalized care delivery model utilizing an interactive Web-based disease management system integrated with the EHR for hypertension management. Compared with baseline, participants significantly reduced office-measured and home-monitored BP at 6 months, with 55.9% of the participants achieving office BP goals, 71.4% achieving home BP goals, and 86.0% achieving clinically meaningful reduction in office BP. The EMPOWER-H program also significantly decreased participants' body weight; increased consumption frequency of fruit and vegetables, minutes of aerobic exercise, and knowledge; and reduced consumption frequency of high-salt and high-fat foods.

The findings of this study are consistent with previous studies showing that technology-assisted clinical tools and approaches hold great promise in improving the quality of hypertension management in the real world. The EMPOWER-H program led to a greater or similar reduction in systolic and diastolic office BP at 6 months compared with the amount of within-group reductions at 6 or 12 months achieved in previous information technology-supported interventions in primary care settings [34-39]. The EMPOWER-H program also resulted in similar percentage of patients meeting office BP goals compared with those achieving BP control in previous interventions [35,37].

The number of intervention contacts showed contrasting associations for office and home BP outcomes. Greater improvement in systolic or diastolic office BP was significantly associated with more intervention contacts, reflecting higher levels of clinical management. In contrast, higher odds of achieving home BP goals were significantly associated with fewer intervention contacts, reflecting lower levels of clinical management. We hypothesize that these results may be because of the clinical management challenges of different subsets of patients. Observation identified that some patients were highly motivated and able to meet their home BP goals independently, using their home BP data and adjustments in personal behavior. Other patients were less motivated, requiring significant clinical attention and more intervention contacts but achieving only moderate improvement in BP as manifested in office BP results. The contrasting results may also be partially explained by changes in the need for intervention contact over time, with more contacts needed in the beginning of the study and less contacts needed once some improvement was seen in home BP values.

Population Health and Patient-Generated Health Data

Technology-assisted interventions in primary care are well aligned with a population health management strategy. About 90% of the American adults with poorly controlled hypertension have health insurance coverage and have received health care in the past year [40], suggesting an important opportunity to control hypertension in primary care settings. The US Preventive Services Task Force recommends that primary care providers offer or refer adults with CVD risk factors to intensive behavioral counseling interventions promoting a healthful diet and physical activity [41]. However, health care providers, including PCPs, nurse practitioners, and RNs usually do not have time or training to deliver lifestyle counseling themselves [42-44], suggesting that patient self-management is important for behavior change and disease management. In addition, the systems necessary to support self-management for larger groups of patient have not been readily available. Information technology-supported interventions in primary care, such as EMPOWER-H, have the potential to achieve hypertension control through facilitating patient engagement with support from health care providers. This study provides lessons for how such systems might be established effectively. The experience gained in this study provides support for the feasibility and value of using patient-generated health data, if properly structured and carefully collected, in the day-to-day clinical management of patients with chronic conditions. It allowed direct insights to clinical progress in the context of day-to-day life that were shared between patient and practitioner but without the need for the patient to visit the office. Reducing the need for office visits can potentially make such population health management strategies more efficient for patients, providers, and the health care system.

However, a number of barriers to wider use exist, and this study provides some insight into how these barriers may be overcome. In this study, whereas the primary outcome measure was office-measured BP, home-monitored values became the primary metric for day-to-day clinical management decisions. Using AHA recommended normal levels for home-monitored BP values provided clinical validity. The fact that patients could not alter the BP measurements allowed practitioners to build trust in the data over time. In addition, regular uploads showing patterns of BP changes, including insights to BP variability by time-of-day (eg, high BP only at night or only in the morning), presented new opportunities for personalized management. Although it is accepted that targets for home-monitored BP are lower than office-measured BP targets [20,45], there is limited real-world experience in the use of these lower targets for setting personal goals for home-monitored BP [46,47]. In addition, practicing clinicians are reported to value home-monitored BP less than office-measured BP for clinical management. They cite concerns about reliable monitors, patient instructions, and result interpretation [48], all of which are addressed by EMPOWER-H procedures.

The EMPOWER-H technology can be categorized as a population health software system. These systems are of increasing importance to health providers and payers to support the delivery of value-based care. A diverse range of information systems have been identified as supporting population health and subcategorization based on differentiation of features [49], and functional objectives can assist in understanding the primary aspect of population health management that individual systems are designed to support. Examples include: (1) systems that provide registry or analytic functions to assist the identification of populations with defined diagnoses and problems or facing specific risks, (2) systems that deliver evidence-based tasks known to drive improved outcomes and targeted to specific subpopulations, and (3) systems that drive the engagement of patients to comply with their personal program of care and make sustained changes to behaviors that positively impact their health outcomes. The EMPOWER-H system falls into this last category.

Limitations

Our study has some methodological limitations such as small sample size, short intervention and follow-up duration, and lack of a control group. Furthermore, this study used a pre-post design, making it hard to discern whether the observed significant improvements were because of the EMPOWER-H intervention or the usual care treatments. Previous randomized controlled trials evaluated the primary care–based interventions that achieved less or similar amount of reduction in systolic and diastolic office BP compared with EMPOWER-H against a usual care condition [34-39]. The significant between-group differences favoring these interventions [34-39] suggest potential efficacy of EMPOWER-H compared with usual care. Despite these limitations, the study contributes to the current literature by testing an interactive Web-based disease management system linked to clinical workflows by interfacing self-monitoring devices (eg, BP monitor and pedometer), EHR, and a study dashboard for BP, weight, and lifestyle behavior management and communication between patients and providers. In addition, the retention rate of this study (99%) was higher than other similar 6-month studies [38,39,50], suggesting that the participants viewed the EMPOWER-H program as beneficial, and the format of the program delivery as acceptable.

Future studies are needed to answer the following questions: (1) What are the long-term effects of this kind of information technology–assisted interventions on clinical outcomes? (2) How could the approaches be further integrated into standard clinic workflows and the EHR, potentially relying more on patient-generated health data for routine care and reducing the need for office visits? (3) How could the approaches be automated (eg, an automated system for behavioral change, disease management, feedback, and risk notification for provider oversight), supporting patients to effectively self-manage their condition with lifestyle changes, to allow cost savings, and further transfer to the whole health care system at a scalable level? (4) How can these methods be applied to other clinical situations (eg, newly diagnosed hypertension and the management of other chronic diseases)? and (5) What is the cost-effectiveness and economic sustainability of such interventions when delivered at large scale?

Conclusions

This study demonstrates that a Web-based system for BP management, with a focus on home BP monitoring driving personalized feedback and care-plan engagement, can be integrated with the EHR and can improve BP among adult patients with poorly controlled hypertension. Furthermore, this study provides insight to the feasibility and value of using patient-generated health data in the day-to-day management of chronic conditions. Questions about generalizability, scalability, and economic sustainability remain, and therefore, a large-scale pragmatic study with a longer follow-up period is warranted. If future studies show that these questions can be addressed, the program has the potential for widespread positive impact through implementation in primary care settings.

Acknowledgments

The authors are indebted to the following individuals for their instrumental contributions to the conduct of the study: Edward Finn, Shauna Hyde, Pragati Kenkare, Linda Klieman, Rakesh Kumar, Dr Monique Lambert, Shaun McElrath, Christina Nabarrete, Elizabeth Silva, Dr Paul Tang, Tonya Waldemeyer, Dr Annie Wong, and Dr Charles Young. They wish to thank the EMPOWER-H participants and their families who made this study possible. This study was funded by The Moore Foundation.

Conflicts of Interest

ME as founder of Ares Health Systems declares a significant financial interest in the Empower system, having acquired ownership rights from Sutter Health on June 13, 2017.

Multimedia Appendix 1

EMPOWER-H educational nugget examples.

[PDF File (Adobe PDF File), 115KB - jmir_v19i9e311_app1.pdf]

Multimedia Appendix 2

Simplified diagram of EMPOWER-H population dashboard design.

[PNG File, 129KB - jmir_v19i9e311_app2.png]

Multimedia Appendix 3

Comprehensive tables of outcome measures used at baseline and 6 months.

[PDF File (Adobe PDF File), 90KB - jmir_v19i9e311_app3.pdf]

References

1. Mozaffarian D, Benjamin EJ, Go AS, Arnett DK, Blaha MJ, Cushman M, American Heart Association Statistics Committee and Stroke Statistics Subcommittee. Heart disease and stroke statistics--2015 update: a report from the American Heart Association. Circulation 2015 Jan 27;131(4):e29-322. [doi: 10.1161/CIR.0000000000000152] [Medline: 25520374]
2. Chobanian AV, Bakris GL, Black HR, Cushman WC, Green LA, Izzo JL, National Heart, Lung, and Blood Institute Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure, National High Blood Pressure Education Program Coordinating Committee. The seventh report of the joint national committee on prevention, detection, evaluation, and treatment of high blood pressure: the JNC 7 report. JAMA 2003 May 21;289(19):2560-2572. [doi: 10.1001/jama.289.19.2560] [Medline: 12748199]
3. Nwankwo T, Yoon S, Burt V, Gu Q. Hypertension among adults in the United States: national health and nutrition examination survey, 2011-2012. NCHS Data Brief 2013 Oct(133):1-8 [FREE Full text] [Medline: 24171916]
4. Ostchega Y, Yoon S, Hughes J, Louis T. Hypertension awareness, treatment, and control--continued disparities in adults: United States, 2005-2006. NCHS Data Brief 2008 Jan(3):1-8 [FREE Full text] [Medline: 19389317]
5. James J. Health Affairs. 2013 Feb 14. Health policy brief: patient engagement URL: http://healthaffairs.org/healthpolicybriefs/brief_pdfs/healthpolicybrief_86.pdf [accessed 2017-08-17] [WebCite Cache ID 6snF7EyxY]
6. Chapman LS, Pelletier KR. Population health management as a strategy for creation of optimal healing environments in worksite and corporate settings. J Altern Complement Med 2004;10(Suppl 1):S127-S140. [Medline: 15630830]
7. Handmaker K, Hart J. 9 steps to effective population health management. Healthc Financ Manage 2015 Apr;69(4):70-76. [Medline: 26665527]
8. Lobach DF, Kawamoto K, Anstrom KJ, Kooy KR, Eisenstein EL, Silvey GM, et al. Proactive population health management in the context of a regional health information exchange using standards-based decision support. AMIA Annu Symp Proc 2007 Oct 11:473-477 [FREE Full text] [Medline: 18693881]
9. Long DA, Perry TL. An application of survival analysis to population health management program evaluation. Am J Health Promot 2007;21(6):529-533. [Medline: 17674641]
10. Logan AG. Transforming hypertension management using mobile health technology for telemonitoring and self-care support. Can J Cardiol 2013 May;29(5):579-585. [doi: 10.1016/j.cjca.2013.02.024] [Medline: 23618506]
11. Chandak A, Joshi A. Self-management of hypertension using technology enabled interventions in primary care settings. Technol Health Care 2015;23(2):119-128. [doi: 10.3233/THC-140886] [Medline: 25515051]
12. Peixoto AJ. Practical aspects of home and ambulatory blood pressure monitoring. Methodist Debakey Cardiovasc J 2015;11(4):214-218 [FREE Full text] [doi: 10.14797/mdcj-11-4-214] [Medline: 27057289]
13. Sivakumaran D, Earle KA. Telemonitoring: use in the management of hypertension. Vasc Health Risk Manag 2014;10:217-224 [FREE Full text] [doi: 10.2147/VHRM.S36749] [Medline: 24748801]
14. Zullig LL, Melnyk SD, Goldstein K, Shaw RJ, Bosworth HB. The role of home blood pressure telemonitoring in managing hypertensive populations. Curr Hypertens Rep 2013 Aug;15(4):346-355 [FREE Full text] [doi: 10.1007/s11906-013-0351-6] [Medline: 23625207]
15. Phenxtoolkit. Measures from the PhenX toolkit version URL: https://www.phenxtoolkit.org/ [accessed 2017-04-03] [WebCite Cache ID 6pU685I4j]
16. Ajzen I. The theory of planned behaviour: reactions and reflections. Psychol Health 2011 Sep;26(9):1113-1127. [doi: 10.1080/08870446.2011.613995] [Medline: 21929476]
17. Janz NK, Becker MH. The health belief model: a decade later. Health Educ Q 1984;11(1):1-47. [Medline: 6392204]
18. Bandura A. Social foundations of thought and action: a social cognitive theory. Englewood Cliffs, NJ: Prentice-Hall; 1986.
19. Tang PC, Overhage JM, Chan AS, Brown NL, Aghighi B, Entwistle MP, et al. Online disease management of diabetes: engaging and motivating patients online with enhanced resources-diabetes (EMPOWER-D), a randomized controlled trial. J Am Med Inform Assoc 2013 May 1;20(3):526-534 [FREE Full text] [doi: 10.1136/amiajnl-2012-001263] [Medline: 23171659]
20. Pickering TG, Hall JE, Appel LJ, Falkner BE, Graves J, Hill MN, et al. Recommendations for blood pressure measurement in humans and experimental animals: part 1: blood pressure measurement in humans: a statement for professionals from the Subcommittee of Professional and Public Education of the American Heart Association Council on High Blood Pressure Research. Circulation 2005 Feb 8;111(5):697-716 [FREE Full text] [doi: 10.1161/01.CIR.0000154900.76284.F6] [Medline: 15699287]
21. Chambers LW, Kaczorowski J, O'Rielly S, Ignagni S, Hearps SJ. Comparison of blood pressure measurements using an automated blood pressure device in community pharmacies and family physicians' offices: a randomized controlled trial. CMAJ Open 2013 Jan;1(1):E37-E42 [FREE Full text] [doi: 10.9778/cmajo.20130005] [Medline: 25077100]
22. Stamler J, Rose G, Stamler R, Elliott P, Dyer A, Marmot M. INTERSALT study findings. Public health and medical care implications. Hypertension 1989 Nov;14(5):570-577 [FREE Full text] [Medline: 2807518]
23. Cook NR, Cohen J, Hebert PR, Taylor JO, Hennekens CH. Implications of small reductions in diastolic blood pressure for primary prevention. Arch Intern Med 1995 Apr 10;155(7):701-709. [Medline: 7695458]
24. McInnes GT. Lowering blood pressure for cardiovascular risk reduction. J Hypertens Suppl 2005 Apr;23(1):S3-S8. [Medline: 15821449]

25. Block G, Gillespie C, Rosenbaum EH, Jenson C. A rapid food screener to assess fat and fruit and vegetable intake. Am J Prev Med 2000 May;18(4):284-288. [Medline: 10788730]
26. NutritionQuest. Questionnaires and screeners URL: https://nutritionquest.com/assessment/list-of-questionnaires-and-screeners/ [accessed 2017-04-05] [WebCite Cache ID 6pU6V8SwV]
27. Patienteducation. Exercise behaviors URL: http://patienteducation.stanford.edu/research/exercise.html [accessed 2017-04-05] [WebCite Cache ID 6pU6oV25X]
28. Selim AJ, Rogers W, Fleishman JA, Qian SX, Fincke BG, Rothendler JA, et al. Updated U.S. population standard for the veterans RAND 12-item health survey (VR-12). Qual Life Res 2009 Feb;18(1):43-52. [doi: 10.1007/s11136-008-9418-2] [Medline: 19051059]
29. Hosonline. The veterans RAND 12 item health survey (VR-12): what it is and how it is used URL: http://www.hosonline.org/globalassets/hos-online/publications/veterans_rand_12_item_health_survey_vr-12_2007.pdf [accessed 2017-04-05] [WebCite Cache ID 6pU71nH5b]
30. Spiro A, Rogers W, Qian S, Kazis L. Hosonline. 2004. Imputing physicalmental summary scores (PCSMCS) for the veterans SF-12 health survey in the context of missing data URL: http://www.hosonline.org/globalassets/hos-online/publications/hos_veterans_12_imputation.pdf [accessed 2017-08-17] [WebCite Cache ID 6snFls6ay]
31. Kazis LE, Miller DR, Skinner KM, Lee A, Ren XS, Clark JA, et al. Patient-reported measures of health: the veterans health study. J Ambul Care Manage 2004;27(1):70-83. [Medline: 14717468]
32. Hibbard JH, Mahoney ER, Stockard J, Tusler M. Development and testing of a short form of the patient activation measure. Health Serv Res 2005 Dec;40(6 Pt 1):1918-1930 [FREE Full text] [doi: 10.1111/j.1475-6773.2005.00438.x] [Medline: 16336556]
33. Hibbard JH, Stockard J, Mahoney ER, Tusler M. Development of the patient activation measure (PAM): conceptualizing and measuring activation in patients and consumers. Health Serv Res 2004 Aug;39(4 Pt 1):1005-1026 [FREE Full text] [doi: 10.1111/j.1475-6773.2004.00269.x] [Medline: 15230939]
34. McManus RJ, Mant J, Bray EP, Holder R, Jones MI, Greenfield S, et al. Telemonitoring and self-management in the control of hypertension (TASMINH2): a randomised controlled trial. Lancet 2010 Jul 17;376(9736):163-172. [doi: 10.1016/S0140-6736(10)60964-6] [Medline: 20619448]
35. Logan AG, Irvine MJ, McIsaac WJ, Tisler A, Rossos PG, Easty A, et al. Effect of home blood pressure telemonitoring with self-care support on uncontrolled systolic hypertension in diabetics. Hypertension 2012 Jul;60(1):51-57 [FREE Full text] [doi: 10.1161/HYPERTENSIONAHA.111.188409] [Medline: 22615116]
36. Rinfret S, Lussier M, Peirce A, Duhamel F, Cossette S, Lalonde L, et al. The impact of a multidisciplinary information technology-supported program on blood pressure control in primary care. Circ Cardiovasc Qual Outcomes 2009 May;2(3):170-177 [FREE Full text] [doi: 10.1161/CIRCOUTCOMES.108.823765] [Medline: 20031834]
37. Green BB, Cook AJ, Ralston JD, Fishman PA, Catz SL, Carlson J, et al. Effectiveness of home blood pressure monitoring, web communication, and pharmacist care on hypertension control: a randomized controlled trial. JAMA 2008 Jun 25;299(24):2857-2867 [FREE Full text] [doi: 10.1001/jama.299.24.2857] [Medline: 18577730]
38. Earle KA, Istepanian RS, Zitouni K, Sungoor A, Tang B. Mobile telemonitoring for achieving tighter targets of blood pressure control in patients with complicated diabetes: a pilot study. Diabetes Technol Ther 2010 Jul;12(7):575-579. [doi: 10.1089/dia.2009.0090] [Medline: 20597833]
39. Magid DJ, Ho PM, Olson KL, Brand DW, Welch LK, Snow KE, et al. A multimodal blood pressure control intervention in 3 healthcare systems. Am J Manag Care 2011 Apr;17(4):e96-103. [Medline: 21774100]
40. Centers for Disease Control and Prevention (CDC). Vital signs: awareness and treatment of uncontrolled hypertension among adults--United States, 2003-2010. MMWR Morb Mortal Wkly Rep 2012 Sep 07;61:703-709 [FREE Full text] [Medline: 22951452]
41. LeFevre ML, U.S. PSTF. Behavioral counseling to promote a healthful diet and physical activity for cardiovascular disease prevention in adults with cardiovascular risk factors: U.S. Preventive Services Task Force Recommendation Statement. Ann Intern Med 2014 Oct 21;161(8):587-593. [doi: 10.7326/M14-1796] [Medline: 25155419]
42. Howe M, Leidel A, Krishnan SM, Weber A, Rubenfire M, Jackson EA. Patient-related diet and exercise counseling: do providers' own lifestyle habits matter? Prev Cardiol 2010;13(4):180-185 [FREE Full text] [doi: 10.1111/j.1751-7141.2010.00079.x] [Medline: 20860642]
43. Jansink R, Braspenning J, van DW, Elwyn G, Grol R. Primary care nurses struggle with lifestyle counseling in diabetes care: a qualitative analysis. BMC Fam Pract 2010 May 25;11:41 [FREE Full text] [doi: 10.1186/1471-2296-11-41] [Medline: 20500841]
44. Morrison F, Shubina M, Turchin A. Lifestyle counseling in routine care and long-term glucose, blood pressure, and cholesterol control in patients with diabetes. Diabetes Care 2012 Feb;35(2):334-341. [doi: 10.2337/dc11-1635] [Medline: 22275442]
45. Pickering T. Recommendations for the use of home (self) and ambulatory blood pressure monitoring. American society of hypertension ad hoc panel. Am J Hypertens 1996 Jan;9(1):1-11. [Medline: 8834700]

46. Jones MI, Greenfield SM, Bray EP, Hobbs FR, Holder R, Little P, et al. Patient self-monitoring of blood pressure and self-titration of medication in primary care: the TASMINH2 trial qualitative study of health professionals' experiences. Br J Gen Pract 2013 Jun;63(611):e378-e385 [FREE Full text] [doi: 10.3399/bjgp13X668168] [Medline: 23735408]
47. Abdullah A, Othman S. The influence of self-owned home blood pressure monitoring (HBPM) on primary care patients with hypertension: a qualitative study. BMC Fam Pract 2011 Dec 30;12:143 [FREE Full text] [doi: 10.1186/1471-2296-12-143] [Medline: 22208768]
48. Logan AG, Dunai A, McIsaac WJ, Irvine MJ, Tisler A. Attitudes of primary care physicians and their patients about home blood pressure monitoring in Ontario. J Hypertens 2008 Mar;26(3):446-452. [doi: 10.1097/HJH.0b013e3282f2fdd4] [Medline: 18300854]
49. PHII. 2016. Population health management software: an opportunity to advance primary care and public health integration URL: http://phii.org/sites/www.phii.org/files/resource/files/Population%20Health%20Management%20Software%20Report.pdf [accessed 2017-08-16] [WebCite Cache ID 6pU7ZEUs7]
50. Madsen LB, Kirkegaard P, Pedersen EB. Blood pressure control during telemonitoring of home blood pressure. A randomized controlled trial during 6 months. Blood Press 2008;17(2):78-86. [doi: 10.1080/08037050801915468] [Medline: 18568696]

Abbreviations

AHA: American Heart Associatione
BMi: body mass index
BP: blood pressure
CVD: cardiovascular disease
DBP: diastolic blood pressure
EHR: electronic health record t
EMPOWER: Engaging and Motivating Patients Online With Enhanced Resources
EMPOWER-D: Engaging and Motivating Patients Online With Enhanced Resources-Diabetes
EMPOWER-H: Engaging and Motivating Patients Online With Enhanced Resources-Hypertension
HRQoL: heart-related quality of life
GED: general education diploma
NCM: nurse care manager
OR: odds ratio
PCP: primary care physician
PAM: Patient Activation Measure
PAMF: Palo Alto Medical Foundation
RD: registered dietitian
RN: registered nurse
SAS: Statistical Analysis System
SBP: systolic blood pressure
SD: standard deviation
VR-12: Veterans
RAND: 12-Item Health Survey

Edited by G Eysenbach; submitted 06.04.17; peer-reviewed by D Hafez, M Plegue; comments to author 11.05.17; revised version received 29.06.17; accepted 27.07.17; published 19.09.17.

Please cite as:
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;19(9):e311
URL: http://www.jmir.org/2017/9/e311/
doi:10.2196/jmir.7831
PMID:28928111

Original Paper

Web-Based Aftercare for Women With Bulimia Nervosa Following Inpatient Treatment: Randomized Controlled Efficacy Trial

Corinna Jacobi[1*], PhD; Ina Beintner[1*], PhD; Eike Fittig[2], PhD; Mickey Trockel[3], MD, PhD; Karsten Braks[4], Dipl. Psych.; Carmen Schade-Brittinger[5], MA; Astrid Dempfle[6], PhD

[1]Institut für Klinische Psychologie und Psychotherapie, Professur Klinische Psychologie & E-Mental Health, Technische Universität Dresden, Dresden, Germany

[2]Celenus Klinik Carolabad, Medizinisches Rehabilitationszentrum für Psychotherapie, Psychiatrie und Psychosomatik, Chemnitz, Germany

[3]Department of Psychiatry and Behavioral Sciences, Stanford University School of Medicine, Stanford, CA, United States

[4]Klinik am Korso gGmbH, Bad Oeynhausen, Germany

[5]Koordinierungszentrum für Klinische Studien Marburg, Philipps-Universität Marburg, Marburg, Germany

[6]Institut für Medizinische Informatik und Statistik, Universitätsklinikum Schleswig-Holstein, Christian-Albrechts-Universität zu Kiel, Kiel, Germany

[*]these authors contributed equally

Corresponding Author:
Corinna Jacobi, PhD
Institut für Klinische Psychologie und Psychotherapie
Professur Klinische Psychologie & E-Mental Health
Technische Universität Dresden
Chemnitzer Strasse 46
Dresden, 01187
Germany
Phone: 49 351 463 38576
Fax: 49 351 463 37208
Email: corinna.jacobi@tu-dresden.de

Abstract

Background: Relapse rates in bulimia nervosa (BN) are high even after successful treatment, but patients often hesitate to take up further treatment. An easily accessible program might help maintain treatment gains. Encouraged by the effects of Web-based eating disorder prevention programs, we developed a manualized, Web-based aftercare program (IN@) for women with BN following inpatient treatment.

Objective: The objective of this study was to determine the efficacy of the web-based guided, 9-month, cognitive-behavioral aftercare program IN@ for women with BN following inpatient treatment.

Methods: We conducted a randomized controlled efficacy trial in 253 women with DSM-IV (Diagnostic and Statistical Manual of Mental Disorders, fourth edition) BN and compared the results of IN@ with treatment as usual (TAU). Assessments were carried out at hospital admission (T0), hospital discharge/baseline (T1), postintervention (T2; 9 months after baseline), 9-month follow-up (T3; 18 months after baseline). The primary outcome, abstinence from binge eating and compensatory behaviors during the 2 months preceding T2, was analyzed by intention to treat, using logistic regression analyses. Frequencies of binge eating and vomiting episodes, and episodes of all compensatory behaviors were analyzed using mixed effects models.

Results: At T2, data from 167 women were available. There were no significant differences in abstinence rates between the TAU group (n=24, 18.9%) and the IN@ group (n=27, 21.4%; odds ratio, OR=1.29; *P*=.44). The frequency of vomiting episodes in the IN@ group was significantly (46%) lower than in the TAU group (*P*=.003). Moderator analyses revealed that both at T2 and T3, women of the intervention group who still reported binge eating and compensatory behaviors after inpatient treatment benefited from IN@, whereas women who were already abstinent after the inpatient treatment did not (*P*=.004; *P*=.002). Additional treatment utilization was high in both groups between baseline and follow-up.

Conclusions: Overall, data from this study suggest moderate effects of IN@. High rates of outpatient treatment utilization after inpatient treatment may have obscured potential intervention effects on abstinence. An aftercare intervention might be more beneficial as part of a stepped-care approach.

Trial Registration: International Standard Randomized Controlled Trial Number (ISRCTN): 08870215; http://www.isrctn.com/ISRCTN08870215 (Archived by WebCite at http://www.webcitation.org/6soA5bIit)

(J Med Internet Res 2017;19(9):e321) doi:10.2196/jmir.7668

KEYWORDS

bulimia nervosa; web-based intervention; aftercare; relapse

Introduction

Bulimia nervosa (BN) is a severe disorder associated with serious medical morbidity and psychosocial comorbidity [1,2]. For a substantial proportion of patients, the long-term course is chronic; depending on outcome criteria and follow-up duration, less than half of the patients achieve full recovery [3,4]. Several studies indicate that higher frequencies of binge eating and purging at baseline, postintervention, and follow-up are negative prognostic factors [5-11].

In the past decades, a number of effective psychological treatments for BN have been developed and are available [12]. Both US (APA) and European guidelines [12,13] recommend cognitive behavioral therapy (CBT [14]) as first-line treatment. Meta-analytic reviews [13,15,16] show good effects for eating disorder attitudes and moderate remission rates ranging between 20% and 30% for CBT. However, relapse rates within the first months of treatment are high even after significant reduction of core BN symptoms [8,17-19], and some women may even be reluctant to seek further treatment after experiencing relapses [20]. Maintaining treatment gains remains a challenge especially for the more severe and chronically ill BN patients [21].

Outside the eating disorders field, technology-enhanced approaches (ie, interventions delivered through the Internet or mobile apps) have been increasingly utilized as aftercare interventions following in- or outpatient treatment across mental health disorders (eg, [22-25]) or for specific disorders such as bipolar disorder [26] or pain disorder [27].

Compared with face-to-face interventions, Web-based interventions offer several potential advantages [28]. They can overcome existing barriers such as cost, service availability, wait time, transportation, and stigma, thus countering health care–related disparities, and can be easily tailored to individual needs. Greater anonymity may encourage individuals to seek help and reveal more sensitive health information. Finally, Web-based programs might be more cost-effective compared with face-to-face interventions [29] and have high user acceptability.

In the field of indicated prevention of eating disorders, Web-based approaches have proven to be efficacious in reducing core symptoms of eating disorders, including high weight and shape concerns, body dissatisfaction [30,31], binge eating [32], new onset of full-syndrome eating disorders [33,34], and shown promise for improving low body weight and restrained eating [35]. However, at the time of initiation of the current trial, no technology-based intervention was available for the maintenance of in- or outpatient treatment gains and prevention of relapses, and in the meantime, very few randomized controlled trials have utilized and tested the efficacy of technology-enhanced interventions following in- or outpatient treatment.

In an aftercare program based on text messaging [36], 165 German patients with BN were randomized to an intervention group (16-week short message service (SMS)–based maintenance) or a treatment as usual (TAU) control group following inpatient treatment. Remission rates at 8 months after hospital discharge were significantly higher in the intervention group than in the TAU group, independent of utilization of outpatient treatment. For patients of the intervention group who did not use any additional outpatient treatment, remission rates were even higher. Another German study [37] examined the efficacy of a Web-based relapse prevention program (RP) over 9 months after inpatient treatment in 258 women with anorexia nervosa (AN), randomized to the RP or TAU condition. At postintervention, RP completers had gained significantly more body weight (0.62 BMI points) than patients in the TAU condition and showed significant improvements in specific eating-related cognitions and behaviors. At 9-month follow-up, the subgroup of participants with high adherence to the program (“full completers”) achieved further improvements in body mass index (BMI) [38]. Finally, following an earlier pilot study with promising results conducted in Hungary, the effects of a Web-based support program following routine outpatient and inpatient care were also evaluated in a recent larger, randomized controlled trial [39]. A total of 105 women with BN and eating disorders not otherwise specified were randomly assigned to an immediate Web-based support program over 4 months or to a 4-month waiting-list TAU control condition. Both groups showed significant reductions in eating disorder-related attitudes at postintervention compared with baseline, with no significant effects of intervention on improvement.

Overall, the results of studies directed at the maintenance of treatment gains for different ED symptoms using technology-based interventions are encouraging, but do not show clear superiority for a relevant outcome measure. These trials also demonstrated that a substantial proportion of patients (28-93%) had utilized additional treatment [36,37,39].

Encouraged by the high acceptance [40] and efficacy of previously evaluated, Web-based indicated prevention programs for eating disorders [30,32,35], we decided to develop an Web-based aftercare program (IN@) for women with BN following inpatient treatment [41]. The aim of this study was to evaluate the efficacy of this program in maintaining or achieving abstinence from core BN symptoms in comparison with a TAU control group. We hypothesized that the aftercare intervention would lead to larger and more stable effects in core BN symptoms compared with TAU.

Methods

Study Design and Participants

We conducted a randomized, controlled efficacy trial in women with DSM-IV BN. The trial protocol, summarizing details on study design and intervention, has been published elsewhere [41]. Over a 4½-year period, patients were screened and recruited from 13 psychosomatic hospitals offering specialized inpatient treatment for eating disorders throughout Germany. Patients were eligible for inclusion if they were aged 17 years or older, fulfilled DSM-IV-TR criteria for BN at hospital admission, and had successfully completed inpatient treatment, defined by a reduction of binge eating and compensatory behaviors by at least 50% based on the past 2 weeks before discharge compared with admission. Patients were excluded if their BMI had dropped below 17.5 kg/m^2 during the inpatient treatment or if local hospital staff regarded them as unfit to participate in a Web-based program (eg, due to psychotic symptoms, acute suicidality, severe personality disorders, or language barriers).

Assessment points were as follows: prebaseline (T0; hospital admission), baseline (T1; hospital discharge), postintervention (T2; 9 months after baseline), and 9-month follow-up (T3; 18 months after baseline). Patients received up to 80€ for participating in all assessments, whereas there were no monetary incentives for using the Web-based intervention.

The study was approved by the local ethics committee of TU Dresden and by ethics committees of all other federal states in which participating hospitals were located. Written informed consent was obtained from all patients and—in case of minors—their legal guardians. The study was conducted according to the Declaration of Helsinki and Good Clinical Practice principles.

Quality-control methods comprised case-report forms, independent data management, on-site monitoring, and documentation of adverse and severe adverse events. Data management was provided by the independent Clinical Trials Center in Marburg and included regular checks for consistency and plausibility, and queries if inconsistencies or missing data became evident.

Randomization and Masking

Concealed randomization was carried out centrally by the independent Clinical Trials Center in Marburg after patients had been enrolled in the study and had given informed consent in a ratio of 1:1. The randomization was done with permuted blocks stratified by Center. Patients and psychologists involved in the moderation of the aftercare program could not be masked to intervention allocation. Assessors who carried out T0-T3 diagnostic assessments were blind to intervention allocation and neither involved in the moderation of the intervention nor in the final data analyses.

Procedures

Within the first 2 weeks after hospital admission, potentially eligible patients were contacted by hospital staff and informed about the study. Patients willing to participate were interviewed over the phone by trained interviewers at TU Dresden to confirm a DSM-IV diagnosis of BN and to assess comorbidity and received login data to access the password protected Web-based trial platform. During the week before or after discharge, patients were interviewed and asked to complete the Web-based assessments again. Patients who had reduced their core BN symptoms by 50% compared with hospital admission were then randomized to the IN@ or TAU condition.

The Intervention IN@

We designed IN@ to target the maintenance of inpatient treatment gains and reduce relapses after discharge. The intervention was based on principles of cognitive behavioral treatment for BN [14] and covered topics such as, eating behaviors and core bulimic symptoms, healthy exercise, body image, self-esteem, emotional and social skills. Table 1 provides an overview of the program content. Additional interactive features of the program were a monitoring log for bulimic symptoms, a body image, and a personal diary. Three clinical psychologists trained in behavior therapy for eating disorders guided the program, that is, provided individualized email feedback to entries in diaries and offered up to 9 monthly real-time individual chats of approximately 1 hour per participant. The program consisted of 11 Web-based sessions over 9 months with fortnightly sessions scheduled during the first 2 months after hospital discharge and monthly sessions thereafter.

The program's home page also provided an overview of all participating hospitals and CVs of the program moderators once patients had logged in to facilitate the credibility of the intervention. Screenshots of the intervention are provided in Multimedia Appendices 1 and 2.

Treatment as Usual (TAU)

Patients assigned to the TAU group were assessed at all assessment points but did not receive additional treatment recommendations from the research team. However, most hospitals recommend some form of outpatient treatment following inpatient treatment. For ethical reasons, we did not interfere with these recommendations and subsequent decisions on concomitant treatment but documented it at all assessment points for both treatment arms.

Table 1. Overview of the program content.

Session	Module	Examples of content
1	Introduction	
	Eating behaviors	Personal eating disorder history
		Dietary restraint and binge eating
		Set-point theory
		Forbidden foods
2	Eating behaviors	Behavioral chain model for identifying high-risk situations
	Body image	Cultural beauty ideals and their impact on self-esteem
	Emotion regulation	Automatic thoughts and their consequences
3	Interpersonal relationships and social skills	Introduction to social skills training
	Emotion regulation	Role of emotions in driving functional and dysfunctional behaviors
4	Interpersonal relationships and social skills	Functional and dysfunctional beliefs and their impact on interpersonal relationships
	Emotion regulation	Role of emotions in triggering dysfunctional eating behaviors
	Eating behaviors	Helpful and unhelpful eating habits Dieting
5	Emotion regulation	Coping with unpleasant emotions
	Alternative behaviors	Increasing pleasant activities
	Body image	Components of body image
		Irrational beliefs regarding body image
6	Interim self-assessment	
7	Emotion regulation	Coping with stress
	Perfectionism	Irrational perfectionist beliefs
8	Body image	Avoidance behavior
	Interpersonal relationships and social skills	Dealing with critical comments
9	Emotion regulation	Mindfulness
	Body image	Mirror confrontation
	Exercise	Reduction of compulsive exercise
10	Self-assessment	
11	Summary	

Outcome Measures

The primary outcome was defined as abstinence from any core BN symptoms (binge eating, vomiting, laxative abuse, abuse of diuretics or other medication to control weight, and driven exercise) in the past 2 months before postintervention (T2) and not having resumed inpatient treatment after T1. Information on the primary outcome was derived both from the Structured Interview for Anorexia and Bulimia nervosa (SIAB-EX) and the weekly symptom checklists. If there was contradictory information from these two sources, the patient was classified as nonabstinent. Secondary outcomes were abstinence from any core BN symptoms in the 2 months before follow-up (T3) as defined earlier, the frequencies of episodes of binge eating, vomiting, and all compensatory behaviors per week in the past 3 months before postintervention (T2) and follow-up (T3), and recovery defined as not fulfilling DSM-IV-TR diagnostic criteria for an eating disorder anymore at postintervention (T2) and follow-up (T3). As abstinence rates are strongly influenced by definitions of abstinence [17], we chose to align our definition of abstinence with previous recommendations [17]. Regarding the frequencies of binge eating and compensatory behaviors as well as recovery, we aligned our definitions with the DSM-IV-TR criteria.

Because hospitals were located nationwide and patients lived in and returned to different parts of Germany, all interview assessments were carried out over the telephone. At all assessment points, trained assessors blinded to patients' group assignment assessed eating disorder-specific and general psychopathology. We applied the SIAB-EX [42] to measure eating disorder symptomatology and general psychopathology and ascertain the diagnosis of BN at all assessment points. The interview was slightly modified by adding continuous items measuring the absolute frequencies (rather than categorical items of frequency spans) of binge eating and compensatory behaviors over periods of 2 weeks, 2 months, and 3 months. At prebaseline, we also used the Structured Clinical Interview for DSM-IV axis I mental disorders (SCID-I [43]) to measure psychiatric comorbidity. All interview documentations were monitored and checked for plausibility by the study coordinator and the independent monitor, and assessors received regular feedback based on this monitoring.

Questionnaire assessments (for further moderator as well as mediator analyses) and the weekly symptom checklists were integrated into the Web-based platform that also hosted the intervention. Patients in both the intervention and the TAU group were prompted regularly to report the frequencies of all core BN symptoms.

At T2 and T3, we also assessed the amount of any additional in- and outpatient treatment patients had utilized during the IN@ intervention and the follow-up period (in-, out-, day-patient therapy, days of treatment, number of outpatient treatment sessions, and number of therapists).

Statistical Analyses

Sample size calculations were based on average rates of patients who do not remain abstinent of binge eating and compensatory behaviors after having achieved abstinence during treatment [17]. We assumed an average abstinence rate of 15% for TAU and of 35% for IN@, that is, we considered a difference of 20 percent points between groups, a clinically relevant reduction in core BN symptomatology. Using standard sample size formulae for the comparison of two rates by Fisher exact test at an alpha-level of 5% (two-sided) and statistical power of 85%, we calculated the minimum sample size to be n=90 for each group or a total of N=180 patients. Assuming a dropout rate of 30% at postintervention, a total of N=258 patients had to be included in the study.

We analyzed the primary outcome by intention to treat (ITT), which included all patients who had been randomized, using logistic regression analyses adjusted for prebaseline binge eating frequency, frequency of compensatory behaviors, BMI, SIAB general psychopathology, and baseline abstinence. If no data (from SIAB-EX or weekly symptom checklist) were available for a patient or if data from both sources were contradictory, she was considered nonabstinent (worst case imputation). The primary outcome was also analyzed using the complete cases (CCs) that included only women who had completed postintervention assessments and the per-protocol sample (PP) which included all women of the IN@ group who had completed postintervention interview assessments and had actively engaged with the intervention, that is, opened at least 25% of program pages or participated in at least two one-to-one chats. For the TAU group, the PP was identical with the CC. Abstinence at follow-up was analyzed in the same manner. Group differences were analyzed by logistic regression analyses, adjusting for prebaseline binge eating frequency, frequency of compensatory behaviors, BMI and SIAB general psychopathology, and baseline abstinence.

We analyzed the frequencies of episodes of binge eating, vomiting, and of all compensatory behaviors per week in the past 3 months preceding baseline (T1, discharge), postintervention (T2), and follow-up (T3) using mixed effects model to account for the nested data structure of three observations across time within individual participants [44]. Longitudinal mixed effects modeling is the method of choice for intent-to-treat analyses to assess clinical trial outcomes [45]. Multiple imputation before conducting longitudinal mixed effects modeling does not add value and is not necessary [46,47]. Episodes of binge eating, vomiting, and other compensatory behaviors were measured as count data and modeled as log-linked dependent variables to estimate intervention effects on change in rate from baseline to T2 or T3.

The percentage of women who did not fulfill diagnostic criteria for any eating disorders at T2 and T3 were compared between the groups by Fisher exact test.

Moderator analyses were performed for abstinence (from binge eating and compensatory behaviors) at postintervention and follow-up using mixed effects model. Because IN@ represents a novel aftercare approach and findings on patient characteristics as predictors of outcome for ED treatment are inconclusive, we decided to employ an exploratory approach as proposed by the MacArthur Foundation [48,49]. Accordingly, in a first step, we calculated correlations (Kendall's Tau) between all available prebaseline and baseline variables and abstinence at postintervention and follow-up. A variable was considered a potential moderator if the correlation between the variable and abstinence at postintervention or follow-up was >.20. In a second step, each candidate moderator variable was entered separately into a mixed effects model. A variable was considered a moderator if there was a significant group × variable interaction effect on either postintervention or follow-up abstinence. Finally, if abstinence was moderated by more than one variable, all moderators were entered in a mixed effects analysis to determine the final prediction model. Statistical analyses were performed using SPSS 22 (IBM, Armonk, NY) and HLM7 (Scientific Software International, Inc., Skokie, IL).

Results

Patient Flow and Characteristics

Between 2007 and 2012, about 2500 patients were admitted for eating disorder treatment to collaborating hospitals and assessed for eligibility. A total of 431 women gave written informed consent to participate in the trial, of which 364 were reached for the prebaseline assessments, and of those, 315 fulfilled DSM-IV-TR criteria for BN (Figure 1).

Figure 1. Study design and recruitment.

Inpatient treatment comprised CBT-oriented (22%) and mixed (psychodynamic with CBT-elements; 78%) treatment approaches. At T1, 274 patients completed the baseline assessment and of those, 253 patients had reduced their core BN symptoms by at least 50% and were thus included in the trial (which would still result in 84% power under the initial planning assumptions); 126 were randomized to IN@ and 127 to TAU. At T2, 88 patients of the intervention group and 79 patients of the control group had completed postintervention assessments resulting in a dropout rate of 30% in the IN@ group and 38% in TAU. At T3, 82 patients of the IN@ group and 81 of the TAU group had completed follow-up assessments resulting in an overall dropout rate of 34.9% in IN@ and 36.2% in TAU.

Table 2 summarizes baseline characteristics of all patients. There were no relevant baseline differences between groups regarding BMI, age, illness duration, frequency of core BN symptoms, and frequency of any prior (in- and outpatient) psychiatric or psychotherapeutic treatment. However, a history of AN was more prevalent in the TAU group (45.6% vs 32.5%), whereas a current anxiety disorder was more prevalent in the IN@ group (46.0% vs 32.3%) and a current substance use disorder was more prevalent in the TAU group (6% vs 1%). Approximately 85.3% of patients had already received some kind of psychiatric or psychotherapeutic treatment before hospital admission.

Overall, patients' BMI was in the normal range. Patients reported illness duration of 7.2 years on average and reported high levels of binge eating and vomiting episodes. In the 2 weeks preceding hospital discharge (baseline), the average number of objective binge eating and vomiting episodes per week had dropped considerably and 54.9% of patients were abstinent from binge eating and compensatory behaviors. Current or lifetime comorbidity was high.

Program Adherence

In total, 19 (15.1%) intervention participants never logged on to the program. Of the remaining 107 women, 46 (36.5%) accessed at least half of the intervention content. On average, participants opened 36% of the program pages (Median: 28%) and accessed 5 of 11 sessions (Median: 4 sessions). Furthermore, 47 women (37.3%) took part in at least one live chat. More detailed information on adherence and therapist time invested will be reported separately.

Effects of the Intervention on Abstinence Rates and Core BN Symptoms

The analysis of the primary outcome showed that at postintervention, about 1 in 5 patients in the ITT sample reported abstinence from binge eating and any compensatory behaviors during the previous 2 months, with no significant difference between the intervention group (21.4%) and the TAU group (18.9%; Table 3). Additional analyses also revealed no significant differences between the groups: In the CC sample (n=167), abstinence rates were 31.0% in the intervention group and 28.2% in the TAU group (*P*=.53). In the PP sample (n=68), the abstinence rate was 30.1% (*P*=.76). Abstinence rates at follow-up (a secondary outcome) remained stable, with no significant difference between the intervention group (22.2%) and the TAU group (17.3%).

Table 2. Sample key characteristics at hospital admission by group.

Patient characteristics	IN@[a], n=126	TAU[b], n=127
Body mass index, in kg/m^2, at prebaseline; mean (SD)[c]	21.49 (2.96)	21.99 (3.85)
Age, in years, at prebaseline; mean (SD)[c]	25.67 (7.18)	26.26 (6.92)
Duration of illness in years at prebaseline, mean (SD)[c]	6.62 (5.59)	7.65 (6.28)
Objective binge eating episodes per week in the past 3 months at prebaseline, mean (SD)[c]	13.72 (11.93)	15.67 (17.11)
Objective binge eating episodes per week in the past 2 weeks at baseline, mean (SD)[c]	0.48 (1.37)	0.66 (2.09)
Vomiting episodes per week in the past 3 months at prebaseline, mean (SD)[c]	18.10 (19.67)	18.73 (20.44)
Vomiting episodes per week in the past 2 weeks at baseline, mean (SD)[c]	0.63 (1.47)	0.80 (2.16)
Episodes of all compensatory behaviors per week in the past 3 months at prebaseline, mean (SD)[c]	22.57 (20.31)	23.39 (20.13)
Episodes of all compensatory behaviors per week in the past 2 weeks at baseline, mean (SD)[c]	1.49 (2.48)	1.71 (2.96)
Abstinent from binge eating and compensatory behaviors in the past 2 weeks at baseline, n (%)[c]	68 (54.0)	71 (55.9)
Any prior psychotherapeutic or psychiatric treatment, n (%)	107 (84.9)	107 (84.2)
History of anorexia nervosa (AN), n (%)[c]	41 (32.5)	58 (45.6)
Current affective disorder at prebaseline, n (%)[d]	50 (39.7)	44 (34.6)
Current anxiety disorder at prebaseline, n (%)[d]	58 (46.0)	41 (32.3)
Current substance abuse disorder at prebaseline, n (%)[d]	1 (0.7)	7 (5.5)

[a] IN@: Web-based aftercare intervention.

[b] TAU: Treatment as usual.

[c] According to modified Structured Interview for Anorexia and Bulimia nervosa (SIAB-EX [42]).

[d] According to Structured Clinical Interview for DSM-IV axis (SCID [43]).

Table 3. Abstinence rates by group at postintervention and follow-up (intent-to-treat analysis).

Outcome of abstinence from binge eating and all compensatory behaviors in the past 2 months[a,e]	IN@[b], n=126 n (%)	TAU[c], n=127 n (%)	Odds ratio (95% CI)[d]	*P* value
At postintervention[f]	27 (21.4)	24 (18.9)	1.29 (0.68-2.44)	.44
At follow-up[f]	28 (22.2)	22 (17.3)	1.49 (0.77-2.86)	.24

[a] Patients who did not provide data were classified as nonabstinent.

[b] IN@: Web-based aftercare intervention.

[c] TAU: Treatment as usual.

[d] Abstinence was coded as 1, nonabstinence as 0 in the model.

[e] No binge eating and compensatory episodes in the past 2 months according to modified structured interview for anorexia and bulimia nervosa (SIAB-EX [42]) and weekly symptom checklist.

[f] Group effect in logistic regression analysis adjusted for prebaseline binge eating frequency, frequency of compensatory behaviors, body mass index (BMI), and SIAB general psychopathology and baseline abstinence.

Table 4. Frequencies of binge eating, vomiting, and all compensatory behaviors: final estimation of fixed effects (population-average model with robust standard errors).

Fixed effect	Coefficient	Standard error	*t* ratio	Event rate ratio (95% CI)	*P* value
Frequency of objective binge eating episodes					
Intercept	−0.56	0.194	−2.872	0.57 (0.390-0.843)	.005
Postintervention slope					
Intercept	2.10	0.214	9.800	8.16 (5.353-12.450)	<.001
IN@ intervention effect	−0.32	0.178	−1.810	0.72 (0.510-1.029)	.07
Follow-up slope					
Intercept	1.98	0.222	8.909	7.23 (4.673-11.216)	<.001
IN@ intervention effect	0.03	0.670	0.040	1.03 (0.274-3.851)	.97
Frequency of vomiting episodes					
Intercept	−0.33	0.161	−2.043	0.72 (0.522-0.992)	.04
Postintervention slope					
Intercept	2.39	0.216	11.070	10.94 (7.147-16.746)	<.001
IN@ intervention effect	−0.61	0.208	−2.951	0.54 (0.359-0.815)	.003
Follow-up slope					
Intercept	2.49	0.241	10.297	12.03 (7.479-19.378)	<.001
IN@ intervention effect	−0.53	0.407	−1.292	0.59 (0.265-1.318)	.20
Frequency of episodes of all compensatory behaviors					
Intercept	0.47	0.107	4.399	1.60 (1.294-1.981)	<.001
Postintervention slope					
Intercept	1.79	0.176	10.167	6.01 (4.246-8.508)	<.001
IN@ intervention effect	−0.36	0.187	−1.910	0.70 (0.483-1.011)	.06
Follow-up slope					
Intercept	1.94	0.184	10.565	6.96 (4.847-9.993)	<.001
IN@ intervention effect	−0.44	0.292	−1.490	0.65 (0.364-1.151)	.14

The frequency of binge eating episodes (Table 4) increased significantly in both groups after hospital discharge. At postintervention, the frequency of binge eating episodes in the intervention group was 28% lower than in the TAU group (3.4 vs 4.7 episodes per week), but the difference was not statistically significant. At follow-up, there was no difference between the groups (4.2 vs 4.1 episodes per week).

The frequency of vomiting episodes (Table 3) also increased significantly in both groups after hospital discharge (baseline).

At postintervention, it was 46% lower in the intervention than in the TAU group (4.3 vs 7.9 episodes per week); this difference was statistically significant. At follow-up, the frequency of vomiting episodes in the intervention group was 41% lower than in the TAU group (5.1 vs 8.7 episodes per week), but the difference was no longer significant. The frequency of episodes of all compensatory behaviors (Table 3) increased significantly in both groups after hospital discharge (baseline). Although at postintervention the frequency of episodes of all compensatory behaviors in the intervention group was 30% lower than in the TAU group (6.8 vs 9.8 episodes per week), the difference was not statistically significant. At follow-up, the frequency of episodes of all compensatory behaviors in the intervention group was 35% lower than in the TAU group (7.2 vs 11.1 episodes per week), but the difference was not statistically significant.

At postintervention, 37 (43.5%) women in the IN@ group and 30 (36.5%) in the TAU group who provided data (CCs set) did not meet diagnostic criteria for any eating disorder anymore; the difference was not significant (*P*=.43); whereas, 25 women (29.4%) in the IN@ group and 37 women (45.1%) in the TAU group met criteria for full syndrome BN, 19 (22.3%) in the IN@ group and 9 (11.0%) in the TAU group met criteria for atypical BN, 1 woman in the IN@ group had crossed over to binge eating disorder and 1 woman in the TAU group had crossed over to binge-eating/purging type AN.

At follow-up, 44 (55.6%) women in the IN@ group and 39 (45.8%) in the TAU group who provided data did not meet diagnostic criteria for any eating disorder anymore; the difference was not significant (*P*=.27). However, 19 women (24.1%) in the IN@ group and 28 women (32.9%) in the TAU group met criteria for full syndrome BN; 14 women (17.7%) in the IN@ group and 17 (20.0%) in the TAU group met criteria for atypical BN; 2 women in the IN@ group had crossed over to binge eating disorder and 1 woman in the TAU group had crossed over to binge-eating/purging type AN.

Moderator Analyses

The following variables were identified as candidate moderators for abstinence at postintervention and follow up: number of diagnostic criteria for BN still present at baseline (hospital discharge), BMI at baseline, abstinence from binge eating and compensatory behaviors, frequency of episodes of binge eating, vomiting and all compensatory behaviors in the 2 weeks preceding baseline, and SIAB subscale scores at baseline.

Mixed model analyses revealed a significant group × variable interaction effect only for abstinence from binge eating both at postintervention and follow-up (Table 5).

Within the subgroup of patients who had failed to achieve abstinence at baseline, women in the intervention group had 4.93 (95% CI 1.90-12.80) times greater odds of abstinence at postintervention and follow-up time points, on average (*P*=.001).

Adverse Events

At the end of the intervention period, 13 women (14.8%) in the IN@ group and 13 women (16.5%) in the TAU group reported clinically relevant suicidal thoughts, with no difference between the two groups (X^2_4=3.8; *P*=.44). One of these women from the TAU group reported suicidal thoughts for the first time at the end of the intervention period; all other women had also reported suicidal thoughts at prebaseline or baseline. Four women (two from each group) reported suicide attempts during the intervention period. Three of those women had comorbid major depression.

Table 5. Moderation of abstinence from binge eating and compensatory behaviors at postintervention and follow-up: final estimation of fixed effects (Population-average model with robust standard errors).

Fixed effect[a]	Coefficient	Standard error	Odds ratio (95% CI)	*P* value
Intercept	0.20	0.126	1.22 (0.948-1.568)	.12
Postintervention slope				
Intercept	−2.43	0.329	0.09 (0.046-0.168)	<.001
Main effect: IN@ intervention	1.24	0.441	3.46 (1.451-8.233)	.005
Main effect: baseline abstinence[b]	2.00	0.419	7.38 (3.230-16.858)	<.001
Interaction effect: intervention X baseline abstinence[b]	−1.71	0.595	0.18 (0.056-0.585)	.004
Follow-up slope				
Intercept	−2.70	0.313	0.07 (0.036-0.124)	<.001
Main effect: IN@ intervention effect	1.61	0.4318	5.02 (2.146-11.742)	<.001
Main effect: baseline abstinence[b]	2.21	0.409	9.08 (4.055-20.349)	<.001
Interaction effect: intervention X baseline abstinence[b]	−1.82	0.595	0.16 (0.050-0.523)	.002

[a] Abstinence was coded as 1, nonabstinence as 0 in the model.

[b] Abstinence from binge eating and compensatory behaviors at baseline (hospital discharge).

Table 6. Utilization of treatment as usual (TAU) by group (two women in each group did not provide data on treatment utilization at follow-up).

Utilization of TAU	IN@[a] N_{post}=85 $N_{follow-up}$=77	TAU[b] N_{post}=82 $N_{follow-up}$=83
Any psychotherapeutic or psychiatric outpatient treatment between baseline and postintervention, n (%)	70 (82.3%)	67 (81.7%)
No. of sessions, mean (SD)	21.7 (23.86)	22.9 (23.91)
Any psychotherapeutic or psychiatric outpatient treatment between postintervention and follow-up, n (%)	55 (71.4%)	61 (73.5%)
No. of sessions, mean (SD)	23.2 (49.71)	25.4 (31.16)
Any psychotherapeutic or psychiatric inpatient treatment between baseline and postintervention, n (%)[c]	6 (7.1%)	2 (2.4%)
Any psychotherapeutic or psychiatric inpatient treatment between postintervention and follow-up, n (%)[c]	9 (11.6%)	7 (8.4%)

[a] IN@: Web-based aftercare intervention.

[b] TAU: Treatment as usual.

[c] Excludes short interventions for suicidal tendencies or substance abuse.

Treatment Seeking

A high proportion of patients of both groups utilized psychotherapeutic or psychiatric outpatient treatment between baseline and postintervention as well as postintervention and 9-month follow-up (Table 6). Rates were almost identical in both groups for both time periods as were number of sessions (IN@: 44.9; TAU: 48.3). A small proportion of patients also utilized further inpatient treatment between hospital discharge and follow-up. Women of the intervention group who did not utilize any further treatment during the intervention period were more likely to terminate the IN@ intervention prematurely (7 out of 14 women [50.0%] opened less than 25% of program pages) than women who received additional CBT (7 out of 30 women [23.3%] opened less than 25% of program pages) or psychodynamic therapy (10 out of 26 women [38.5%] opened less than 25% of program pages). However, the type (CBT vs psychodynamic) or amount of TAU treatment was not significantly related to any of the reported outcomes.

Discussion

Principal Findings

This is one of the first studies to evaluate the efficacy of a program targeting maintenance or improvement of treatment gains achieved during inpatient treatment in patients with severe BN. On the basis of ITT analysis at T2, we found that individuals receiving IN@ did not differ significantly from individuals receiving TAU in rates of abstinence from binge eating and any compensatory behaviors (21.4% vs 18.9%) and these rates remained stable at follow-up. They were higher in the CCs (and PP) sample but also not significantly different between groups. Although the intervention resulted in a 28% lower frequency of binge eating episodes at posttreatment compared with the TAU group, this difference was not significant. However, at postintervention, the IN@ group achieved a 46% lower frequency of vomiting episodes and this difference was significant. At follow up, this frequency was still 41% lower than in the TAU group but the difference was no longer significant. Moderator analyses revealed that both at post-intervention and follow-up, women of the intervention group who still reported binge eating and compensatory behaviors after the inpatient treatment had higher odds of achieving abstinence at post-intervention and follow-up compared with women of the TAU group who still reported binge eating and compensatory behaviors after the inpatient treatment.

Comparison With Prior Work

The results of this study have to be discussed in the light of several important points: (1) general lack of comparison of maintenance studies for patients with BN, (2) illness severity, (3) additional treatment utilization and health care system in general, and (4) intervention characteristics.

For the primary outcome, abstinence of binge eating and any compensatory behaviors, only one other study for patients with BN has demonstrated that an aftercare intervention yields a significant reduction in abstinence while one other study did not find any differences on a number of self-report measures between the groups. In the first study, a 16-week, SMS for patients with BN after hospital discharge was investigated [36]. This relatively short intervention yielded significantly larger abstinence and partial remission rates at 8 months after hospital discharge than our intervention but the SMS-study differs in terms of duration of intervention and follow up, diagnoses (more EDNOS cases), and symptom frequency from ours. Additionally, treatment utilization in this study was 8-9 sessions over the 8 months of the intervention and follow-up compared with 27-29 sessions over 9 months of the intervention in our study. Because this study [36] was carried out as part of the same health care system as our own study, patients were probably not as severely ill as patients included in our study. In the second study [39] published after the starting of our study, both the Web-based support group and the waiting list group improved during the 4-month intervention with no significant differences between groups. Neither abstinence rates nor follow-up data were reported.

A second important point for the interpretation of the results of this study regards illness severity of the included patient group, indicated by symptom frequency, illness duration, and previous in- and outpatient treatment utilization. Patients in our study reported an average frequency of almost 15 binge eating episodes and over 18 vomiting episodes per week during the preceding 3 months before hospital admission. Although patients treated as inpatients may have higher frequencies of binge eating and purging episodes in general, these numbers clearly exceed frequencies reported in both earlier and more recent intervention trials for BN (eg, [50-52]). Our sample may therefore represent a more severely ill sample of BN patients which may limit the generalizability of the results.

The third crucial issue to be considered in the interpretation of the results is the amount of additional treatment utilization. Our study was originally planned to compare the efficacy of an aftercare intervention with TAU. However, at the time of planning the study, it was unclear what the "usual" amount of further treatment utilized by patients after hospital discharge would be, that is, whether and to what degree patients are recommended to take up further treatment following their inpatient stay and to what degree they would follow this recommendation. In this study, treatment utilization was surprisingly high: 82% of both groups engaged in additional treatment with 2-3 sessions per month during the 9-month intervention period and even higher numbers over the 9-month follow-up period. For patients, as severely ill as in this study, a Web-based aftercare intervention may have been less acceptable as stand-alone treatment compared with face-to-face treatment.

Finally, treatment utilization in this study may also be reflective of the specific health care context in which the study was conducted, that is, access to outpatient psychotherapeutic care and coverage of up to 80 sessions of outpatient psychotherapeutic treatment by health insurances in Germany. However, waiting times can be 6-12 months depending on regional differences. Because patients originated from different parts in Germany, the amount of further treatment utilization is even more surprising.

The characteristics of the intervention itself may also need to be discussed for the interpretation of the results. Adherence is a problem for Web-based interventions in general, also labeled as "law of attrition" [53]. Compared with targeted preventive interventions for eating disorders in student populations, in which adherence, defined as the number of opened pages and/or completed sessions, ranges between 50% and 80% [40], adherence to the IN@ intervention was lower. Participants opened on average 36% of the program pages; 15.1% of patients never logged on to the program after randomization and only 19.1% completed all 11 sessions of the program. However, the duration of preventive interventions is usually 8-10 weeks compared with 9 months in our study. For the included patients, the frequency of provided Web-based sessions especially during the first months may have been too low and the duration of the intervention and of individual sessions may have been too long. Future maintenance interventions may benefit from an overall shorter intervention duration and shorter, more frequent sessions which may also improve adherence to the intervention.

Clinical and Service-Related Implications

Although the intervention did not have a significant effect on abstinence rates at T2, the results of secondary outcomes, specifically vomiting rates, and moderator analyses may have important clinical implications: In the IN@ group, rates of vomiting episodes were almost half the rates of patients in the TAU group. From a clinical perspective, reducing vomiting episodes is often more difficult for patients than reducing binge eating, and while patients may have given up or reduced the frequency and quality of binge eating episodes considerably, vomiting is often the more persistent symptom. Because vomiting is also specifically associated with severe medical complications, its reduction during treatment is of essential clinical importance [54]. However, despite the marked reductions in vomiting rates during the intervention and follow-up, average rates were still above the diagnostically relevant threshold.

The intervention also proved to be specifically beneficial for patients who had not achieved abstinence at the time of hospital discharge. Thus, for a subgroup of potentially even more severely ill patients, participating in the aftercare intervention turned out to be specifically advantageous. In our study, only 54.9% of patients achieved abstinence of all BN symptoms during inpatient treatment in the past 2 weeks before baseline. Accordingly, a considerable proportion of patients were in need of further or more specific support to achieve full remission. Generally, the lack of studies examining short- and long-term effects of inpatient or residential treatments limits the comparability of these outcomes [55]. However, studies addressing long-term course and outcome of BN often demonstrate that relapses after the end of treatment are common and that patient status at posttreatment does not necessarily predict recovery status at follow-up [56]. The fact that patients in this study did not maintain improvements achieved during inpatient treatment and that a large proportion of patients did not improve markedly despite further treatment utilization and participation in IN@ may also demonstrate the need for a stepped—and more consistent and specialized—aftercare treatment provision. TAU after inpatient treatment is obviously provided irrespective of previous treatment gains or further needs in specific symptom domains. Although for some patients the aftercare intervention may have been sufficient to maintain their treatment gains, others may have needed even more intensive outpatient face-to-face treatment. To answer the question of who might need which amount of further Web-based or face-to-face treatment would require a change in current provision of health care delivery based on evidence rather than nonevidence based exploitation of existing health care structures.

Strengths and Limitations

Our study includes one of the largest sample sizes of patients of intervention trials for patients with BN, which clearly is a strength of the study. Follow-up duration is relatively long, BN symptoms and diagnoses were ascertained by a well-validated clinical interview, and the number of adverse events was not higher in IN@ compared with TAU. In a stepped-care model, the intervention could represent an economic alternative to face-to-face treatment if provided in a shortened version as first step and full alternative to TAU.

Limitations of the study are moderately higher dropout rates compared with face-to-face intervention trials and low adherence to the Web-based intervention. Future research efforts should therefore also be directed at adapting the intervention to increase adherence to the program and reduce dropouts.

Conclusions

Taken together, data from this study suggest moderate effects of IN@ for patients with severe and chronic BN. With regard to abstinence rates, IN@ did not show a marked superiority over TAU. However, the intervention had encouraging effects on vomiting rates and might also be particularly beneficial for women who have not achieved abstinence during previous inpatient treatment when compared with TAU.

Acknowledgments

The study was funded by the German Federal Ministry of Education and Research (Bundesministerium für Bildung und Forschung, BMBF), as part of the German Eating Disorders Diagnostic and Treatment Network (EDNET). (01GV0603).

We thank all patients who participated in this study for their contributions. Thanks to Student Research Assistants: Katharina Möbius, Sina Rübesamen, Anne Ueck; Data Monitoring: Frauke Schmidt; Data Management: Maik Hahmann; Statistical Advice: Nina Timmesfeld; Hospital Staff: Sylvia Beisel, Ute Biersack, Nina Dziuk, Volker Bracke, Ulrike Eichler, Monika Göttel-Leypold, Nadine Hamann, Tina Käshammer, Silke Kleinschmidt, Petra Künker, Lisette Morris, Johannes Schaller, Marthe Schönewald, Petra Stadtfeld-Oertel, Ursula Vatterodt, Martin Wiesjahn, Nadine Ziegler, and Dino Zitarosa; and Assessors: Christina Kampisiou, Anna Dill, Katrin Dittmer, Eileen Eismann, Scarlett Hentzschel, Andrea Horn, Katja Hornemann, Beate Krüger, Heike Melzer, Katharina Möbius, Jana Mrose, Ricarda Ohlmer, Carolin Schreiber, Jana Seltmann, Stephanie Teubert, Dagmar Uhlmann, Alexander Varchmin, Ulrike Völker, and Miriam Wolf

Conflicts of Interest

None declared.

Multimedia Appendix 1

IN@ screenshot 1.

[JPG File, 105KB - jmir_v19i9e321_app1.JPG]

Multimedia Appendix 2

IN@ screenshot 2.

[JPG File, 129KB - jmir_v19i9e321_app2.JPG]

Multimedia Appendix 3

CONSORT-EHEALTH checklist (v1.6.1).

[PDF File (Adobe PDF File), 630KB - jmir_v19i9e321_app3.pdf]

References

1. Hudson JI, Hiripi E, Pope HG, Kessler RC. The prevalence and correlates of eating disorders in the National comorbidity survey replication. Biol Psychiatry 2007 Feb 01;61(3):348-358 [FREE Full text] [doi: 10.1016/j.biopsych.2006.03.040] [Medline: 16815322]
2. Swanson SA, Crow SJ, Le Grange D, Swendsen J, Merikangas KR. Prevalence and correlates of eating disorders in adolescents. Results from the national comorbidity survey replication adolescent supplement. Arch Gen Psychiatry 2011 Jul;68(7):714-723. [doi: 10.1001/archgenpsychiatry.2011.22] [Medline: 21383252]
3. Steinhausen H, Weber S. The outcome of bulimia nervosa: findings from one-quarter century of research. Am J Psychiatry 2009 Dec;166(12):1331-1341. [doi: 10.1176/appi.ajp.2009.09040582] [Medline: 19884225]
4. Castellini G, Lo SC, Mannucci E, Ravaldi C, Rotella CM, Faravelli C, et al. Diagnostic crossover and outcome predictors in eating disorders according to DSM-IV and DSM-V proposed criteria: a 6-year follow-up study. Psychosom Med 2011 Apr;73(3):270-279. [doi: 10.1097/PSY.0b013e31820a1838] [Medline: 21257978]
5. Bulik CM, Sullivan PF, Joyce PR, Carter FA, McIntosh VV. Predictors of 1-year treatment outcome in bulimia nervosa. Compr Psychiatry 1998;39(4):206-214. [Medline: 9675505]
6. Turnbull SJ, Schmidt U, Troop NA, Tiller J, Todd G, Treasure JL. Predictors of outcome for two treatments for bulimia nervosa: short and long-term. Int J Eat Disord 1997 Jan;21(1):17-22. [Medline: 8986513]
7. Fahy TA, Russell GF. Outcome and prognostic variables in bulimia nervosa. Int J Eat Disord 1993 Sep;14(2):135-145. [Medline: 8401546]

8. Olmsted MP, Kaplan AS, Rockert W. Rate and prediction of relapse in bulimia nervosa. Am J Psychiatry 1994 May;151(5):738-743. [doi: 10.1176/ajp.151.5.738] [Medline: 8166317]
9. Helverskov JL, Clausen L, Mors O, Frydenberg M, Thomsen PH, Rokkedal K. Trans-diagnostic outcome of eating disorders: a 30-month follow-up study of 629 patients. Eur Eat Disord Rev 2010;18(6):453-463. [doi: 10.1002/erv.1025] [Medline: 20593480]
10. Marrone S, Mitchell JE, Crosby R, Wonderlich S, Jollie-Trottier T. Predictors of response to cognitive behavioral treatment for bulimia nervosa delivered via telemedicine versus face-to-face. Int J Eat Disord 2009 Apr;42(3):222-227. [doi: 10.1002/eat.20603] [Medline: 18951452]
11. Wilson GT, Loeb KL, Walsh BT, Labouvie E, Petkova E, Liu X, et al. Psychological versus pharmacological treatments of bulimia nervosa: predictors and processes of change. J Consult Clin Psychol 1999 Aug;67(4):451-459. [Medline: 10450615]
12. National Institute for Health and Care Excellence. Eating disorders: Core interventions in the treatment and management of anorexia nervosa, bulimia nervosa and related eating disorders. Leister: British Psychological Society; 2004.
13. Herpertz S, Herpertz-Dahlmann B, Fichter M, Tuschen-Caffier B, Zeeck A. S3-Leitlinie Diagnostik und Behandlung der Essstörungen. Berlin: Springer; 2011.
14. Fairburn C, Marcus M, Wilson G. Cognitive-behavioral therapy for binge eating and bulimia nervosa: a comprehensive treatment manual. In: Binge eating: Nature, assessment, and treatment. New York: The Guildford Press; 1993:361-404.
15. Hay P. A systematic review of evidence for psychological treatments in eating disorders: 2005-2012. Int J Eat Disord 2013 Jul;46(5):462-469. [doi: 10.1002/eat.22103] [Medline: 23658093]
16. Hay PP, Bacaltchuk J, Stefano S, Kashyap P. Psychological treatments for bulimia nervosa and binging. Cochrane Database Syst Rev 2009 Oct 07(4):CD000562. [doi: 10.1002/14651858.CD000562.pub3] [Medline: 19821271]
17. Olmsted MP, Kaplan AS, Rockert W. Defining remission and relapse in bulimia nervosa. Int J Eat Disord 2005 Jul;38(1):1-6. [doi: 10.1002/eat.20144] [Medline: 15971233]
18. Richard M, Bauer S, Kordy H. Relapse in anorexia and bulimia nervosa—a 2.5-year follow-up study. Eur Eat Disorders Rev 2005 May;13(3):180-190. [doi: 10.1002/erv.638]
19. Fichter MM, Quadflieg N, Hedlund S. Long-term course of binge eating disorder and bulimia nervosa: relevance for nosology and diagnostic criteria. Int J Eat Disord 2008 Nov;41(7):577-586. [doi: 10.1002/eat.20539] [Medline: 18473335]
20. Mitchell JE, Agras WS, Wilson GT, Halmi K, Kraemer H, Crow S. A trial of a relapse prevention strategy in women with bulimia nervosa who respond to cognitive-behavior therapy. Int J Eat Disord 2004 May;35(4):549-555. [doi: 10.1002/eat.10265] [Medline: 15101070]
21. Wonderlich S, Mitchell JE, Crosby RD, Myers TC, Kadlec K, Lahaise K, et al. Minimizing and treating chronicity in the eating disorders: a clinical overview. Int J Eat Disord 2012 May;45(4):467-475. [doi: 10.1002/eat.20978] [Medline: 22271525]
22. Ebert DD, Gollwitzer M, Riper H, Cuijpers P, Baumeister H, Berking M. For whom does it work? moderators of outcome on the effect of a transdiagnostic internet-based maintenance treatment after inpatient psychotherapy: randomized controlled trial. J Med Internet Res 2013 Aug;15(10):e191 [FREE Full text] [doi: 10.2196/jmir.2511] [Medline: 24113764]
23. Ebert D, Tarnowski T, Gollwitzer M, Sieland B, Berking M. A transdiagnostic internet-based maintenance treatment enhances the stability of outcome after inpatient cognitive behavioral therapy: a randomized controlled trial. Psychother Psychosom 2013;82(4):246-256. [doi: 10.1159/000345967] [Medline: 23736751]
24. Golkaramnay V, Bauer S, Haug S, Wolf M, Kordy H. The exploration of the effectiveness of group therapy through an internet chat as aftercare: a controlled naturalistic study. Psychother Psychosom 2007;76(4):219-225. [doi: 10.1159/000101500] [Medline: 17570960]
25. Golkaramnay V, Wangemann T, Dogs J, Dogs P, Kordy H. [New bridges for gaps in psychotherapeutic service provision by the internet: hopes, challenges and a solution]. Psychother Psychosom Med Psychol 2003;53(9-10):399-405. [doi: 10.1055/s-2003-42175] [Medline: 14528409]
26. Barnes C, Harvey R, Mitchell P, Smith M, Wilhelm K. Evaluation of an online relapse prevention program for bipolar disorder - an overview of the aims and methodology of a randomized controlled trial. Disease Management & Health Outcomes 2007;15(4):215-224. [doi: 10.1056/NEJM196803212781204]
27. Moessner M, Aufdermauer N, Baier C, Göbel H, Kuhnt O, Neubauer E, et al. [Efficacy of an internet-delivered aftercare program for patients with chronic back pain]. Psychother Psychosom Med Psychol 2014 Feb;64(2):47-53. [doi: 10.1055/s-0033-1351266] [Medline: 24101036]
28. Christensen H, Hickie IB. E-mental health: a new era in delivery of mental health services. Med J Aust 2010 Jun 07;192(11 Suppl):S2-S3. [Medline: 20528702]
29. Portnoy DB, Scott-Sheldon LAJ, 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]
30. Beintner I, Jacobi C, Taylor CB. Effects of an Internet-based prevention programme for eating disorders in the USA and Germany--a meta-analytic review. Eur Eat Disord Rev 2012 Jan;20(1):1-8. [doi: 10.1002/erv.1130] [Medline: 21796737]

31. Stice E, Shaw H, Marti CN. A meta-analytic review of eating disorder prevention programs: encouraging findings. Annu Rev Clin Psychol 2007;3:207-231. [doi: 10.1146/annurev.clinpsy.3.022806.091447] [Medline: 17716054]
32. Jacobi C, Völker U, Trockel MT, Taylor CB. Effects of an internet-based intervention for subthreshold eating disorders: a randomized controlled trial. Behav Res Ther 2012 Feb;50(2):93-99. [doi: 10.1016/j.brat.2011.09.013] [Medline: 22137366]
33. Taylor CB, Bryson S, Luce KH, Cunning D, Doyle AC, Abascal LB, et al. Prevention of eating disorders in at-risk college-age women. Arch Gen Psychiatry 2006 Aug;63(8):881-888 [FREE Full text] [doi: 10.1001/archpsyc.63.8.881] [Medline: 16894064]
34. Stice E, Shaw H, Becker CB, Rohde P. Dissonance-based interventions for the prevention of eating disorders: using persuasion principles to promote health. Prev Sci 2008 Jun;9(2):114-128 [FREE Full text] [doi: 10.1007/s11121-008-0093-x] [Medline: 18506621]
35. Ohlmer R, Jacobi C, Taylor CB. Preventing symptom progression in women at risk for AN: results of a pilot study. Eur Eat Disord Rev 2013 Jul;21(4):323-329. [doi: 10.1002/erv.2225] [Medline: 23520152]
36. Bauer S, Okon E, Meermann R, Kordy H. Technology-enhanced maintenance of treatment gains in eating disorders: efficacy of an intervention delivered via text messaging. J Consult Clin Psychol 2012 Aug;80(4):700-706. [doi: 10.1037/a0028030] [Medline: 22545736]
37. Fichter MM, Quadflieg N, Nisslmüller K, Lindner S, Osen B, Huber T, et al. Does internet-based prevention reduce the risk of relapse for anorexia nervosa? Behav Res Ther 2012 Mar;50(3):180-190. [doi: 10.1016/j.brat.2011.12.003] [Medline: 22317754]
38. Fichter MM, Quadflieg N, Lindner S. Internet-based relapse prevention for anorexia nervosa: nine- month follow-up. J Eat Disord 2013;1:23 [FREE Full text] [doi: 10.1186/2050-2974-1-23] [Medline: 24999404]
39. Gulec H, Moessner M, Túry F, Fiedler P, Mezei A, Bauer S. A randomized controlled trial of an internet-based posttreatment care for patients with eating disorders. Telemed J E Health 2014 Oct;20(10):916-922. [doi: 10.1089/tmj.2013.0353] [Medline: 25188398]
40. Beintner I, Jacobi C, Taylor CB. Participant adherence to the internet-based prevention program student bodies? for eating disorders - a review. Internet Interv 2014;1(1):26-32 [FREE Full text] [doi: 10.1002/iub.1366] [Medline: 25904163]
41. Beintner I, Jacobi C. Internetgestützte Nachsorge bei Bulimia Nervosa. Psychotherapeut 2011 Oct 20;56(6):516-521. [doi: 10.1007/s00278-011-0869-5]
42. Fichter MM, Herpertz S, Quadflieg N, Herpertz-Dahlmann B. Structured interview for anorexic and bulimic disorders for DSM-IV and ICD-10: updated (third) revision. Int J Eat Disord 1998 Nov;24(3):227-249. [Medline: 9741034]
43. Wittchen H, Zaudig M, Fydrich T. SKID: Strukturiertes Klinisches Interview für DSM-IV: Achse I und II. Göttingen: Hogrefe; 1997.
44. Raudenbush S, Bryk A. Hierarchical Linear Models: Applications and Data Analysis Methods. Thousand Oaks, CA: Sage Publications Inc; 2002.
45. Chakraborty H, Gu H. RTI. Research Triangle Park, NC: RTI Press; 2009. A mixed model approach for intent-to-treat analysis in longitudinal clinical trials with missing values URL: https://www.rti.org/sites/default/files/resources/13647477_mr-0009-0904-chakraborty.pdf [accessed 2017-08-29] [WebCite Cache ID 6t4hUDCel]
46. Peters SAE, Bots ML, den Ruijter HM, Palmer MK, Grobbee DE, Crouse JR, et al. Multiple imputation of missing repeated outcome measurements did not add to linear mixed-effects models. J Clin Epidemiol 2012 Jun;65(6):686-695. [doi: 10.1016/j.jclinepi.2011.11.012] [Medline: 22459429]
47. Twisk J, de Boer M, de Vente M, Heymans M. Multiple imputation of missing values was not necessary before performing a longitudinal mixed-model analysis. J Clin Epidemiol 2013 Sep;66(9):1022-1028. [doi: 10.1016/j.jclinepi.2013.03.017] [Medline: 23790725]
48. Kraemer HC, Stice E, Kazdin A, Offord D, Kupfer D. How do risk factors work together? mediators, moderators, and independent, overlapping, and proxy risk factors. Am J Psychiatry 2001 Jun;158(6):848-856. [doi: 10.1176/appi.ajp.158.6.848] [Medline: 11384888]
49. Kraemer HC, Wilson GT, Fairburn CG, Agras WS. Mediators and moderators of treatment effects in randomized clinical trials. Arch Gen Psychiatry 2002 Oct;59(10):877-883. [Medline: 12365874]
50. Agras WS, Walsh T, Fairburn CG, Wilson GT, Kraemer HC. A multicenter comparison of cognitive-behavioral therapy and interpersonal psychotherapy for bulimia nervosa. Arch Gen Psychiatry 2000 May;57(5):459-466. [Medline: 10807486]
51. Poulsen S, Lunn S, Daniel SI, Folke S, Mathiesen BB, Katznelson H, et al. A randomized controlled trial of psychoanalytic psychotherapy or cognitive-behavioral therapy for bulimia nervosa. Am J Psychiatry 2014 Jan;171(1):109-116. [doi: 10.1176/appi.ajp.2013.12121511] [Medline: 24275909]
52. Mitchell JE, Agras S, Crow S, Halmi K, Fairburn CG, Bryson S, et al. Stepped care and cognitive-behavioural therapy for bulimia nervosa: randomised trial. Br J Psychiatry 2011 May;198(5):391-397 [FREE Full text] [doi: 10.1192/bjp.bp.110.082172] [Medline: 21415046]
53. Eysenbach G. The law of attrition. J Med Internet Res 2005;7(1):e11 [FREE Full text] [doi: 10.2196/jmir.7.1.e11] [Medline: 15829473]
54. Mehler PS, Rylander M. Bulimia Nervosa - medical complications. J Eat Disord 2015;3:12 [FREE Full text] [doi: 10.1186/s40337-015-0044-4] [Medline: 25866627]

55. Friedman K, Ramirez AL, Murray SB, Anderson LK, Cusack A, Boutelle KN, et al. A narrative review of outcome studies for residential and partial hospital-based treatment of eating disorders. Eur Eat Disord Rev 2016 Jul;24(4):263-276. [doi: 10.1002/erv.2449] [Medline: 27062687]
56. Lock J, Agras WS, Le Grange D, Couturier J, Safer D, Bryson SW. Do end of treatment assessments predict outcome at follow-up in eating disorders? Int J Eat Disord 2013 Dec;46(8):771-778. [doi: 10.1002/eat.22175] [Medline: 23946139]

Abbreviations

AN: anorexia nervosa
BN: Bulimia nervosa
BMI: body mass index
CBT: cognitive behavioral therapy
DSM-IV: Diagnostic and Statistical Manual of Mental Disorders
RP: relapse prevention
SMS: short message service
TAU: treatment as usual

Edited by G Eysenbach; submitted 10.03.17; peer-reviewed by S Schlegl, U Schmidt; comments to author 04.05.17; revised version received 11.06.17; accepted 04.07.17; published 22.09.17.

Please cite as:
Jacobi C, Beintner I, Fittig E, Trockel M, Braks K, Schade-Brittinger C, Dempfle A
Web-Based Aftercare for Women With Bulimia Nervosa Following Inpatient Treatment: Randomized Controlled Efficacy Trial
J Med Internet Res 2017;19(9):e321
URL: http://www.jmir.org/2017/9/e321/
doi:10.2196/jmir.7668
PMID:28939544

Original Paper

Photoaging Mobile Apps in School-Based Melanoma Prevention: Pilot Study

Titus Josef Brinker[1,2,3], MD; Christian Martin Brieske[1,3,4]; Christoph Matthias Schaefer[1,3,4]; Fabian Buslaff[2,3]; Martina Gatzka[5], MD; Maximilian Philip Petri[1,3,4], MD; Wiebke Sondermann[1,3,4], MD; Dirk Schadendorf[1,3,4], MD, PhD; Ingo Stoffels[1,3,4], MD, PhD; Joachim Klode[1,3,4], MD, PhD

[1]Department of Dermatology, Venereology and Allergology, University-Hospital Essen, University of Duisburg-Essen, Essen, Germany

[2]Department of Dermatology and National Center for Tumor Diseases (NCT), University Hospital Heidelberg, University of Heidelberg, Heidelberg, Germany

[3]German Cancer Consortium (DKTK), Heidelberg, Germany

[4]West German Cancer Center, University Duisburg-Essen, Essen, Germany

[5]Department of Dermatology and Allergic Diseases, University of Ulm, Ulm, Germany

Corresponding Author:
Titus Josef Brinker, MD
Department of Dermatology, Venereology and Allergology
University-Hospital Essen
University of Duisburg-Essen
Hufelandstrasse 55
Essen, 45147
Germany
Phone: 49 +4915175984347
Email: titus.brinker@gmail.com

Abstract

Background: Around 90% of melanomas are caused by exposure to ultraviolet (UV) radiation and are therefore eminently preventable. Tanning behavior is mostly initiated in early adolescence, often with the belief that it increases attractiveness; the problems related to malignant melanoma and other skin cancers are too far in the future to fathom. Given the substantial amount of time children and adolescents spend in schools, as well as with their mobile phones, addressing melanoma prevention via both of these ways is crucial. However, no school-based intervention using mobile apps has been evaluated to date. We recently released a photoaging mobile app, in which a selfie is altered to predict future appearance dependent on UV protection behavior and skin type.

Objective: In this pilot study, we aimed to use mobile phone technology to improve school-based melanoma prevention and measure its preliminary success in different subgroups of students with regard to their UV protection behavior, Fitzpatrick skin type and age.

Methods: We implemented a free photoaging mobile phone app (Sunface) in 2 German secondary schools via a method called mirroring. We “mirrored” the students’ altered 3-dimensional (3D) selfies reacting to touch on mobile phones or tablets via a projector in front of their whole grade. Using an anonymous questionnaire capturing sociodemographic data as well as risk factors for melanoma we then measured their perceptions of the intervention on a 5-point Likert scale among 205 students of both sexes aged 13-19 years (median 15 years).

Results: We measured more than 60% agreement in both items that measured motivation to reduce UV exposure and only 12.5% disagreement: 126 (63.0%) agreed or strongly agreed that their 3D selfie motivated them to avoid using a tanning bed, and 124 (61.7%) to increase use of sun protection. However, only 25 (12.5%) disagreed with both items. The perceived effect on motivation was increased in participants with Fitzpatrick skin types 1-2 in both tanning bed avoidance (n=74, 71.8% agreement in skin types 1-2 vs n=50, 53.8% agreement in skin types 3-6) and increased use of sun protection (n=70, 68.0% agreement in skin types 1-2 vs n=52, 55.3% agreement in skin types 3-6), and also positively correlated with higher age.

Conclusions: We present a novel way of integrating photoaging in school-based melanoma prevention that affects the students’ peer group, considers the predictors of UV exposure in accordance with the theory of planned behavior, and is particularly effective in changing behavioral predictors in fair-skinned adolescents (Fitzpatrick skin types 1-2). Further research is required to evaluate the intervention's prospective effects on adolescents of various cultural backgrounds.

(J Med Internet Res 2017;19(9):e319) doi:10.2196/jmir.8661

KEYWORDS

melanoma; skin cancer; prevention; mobile apps; smartphones; photoaging; schools; secondary schools; adolescents

Introduction

Skin cancer is the most common malignancy in fair-skinned populations, with melanoma incidence being the highest in New Zealand and Australia (50 and 48 per 100,000 population, respectively) and projected to increase in European countries such as the United Kingdom (from 17 to 36 per 100,000 population) between 2007-2011 and 2022-2026 [1].

Around 90% of melanomas are caused by ultraviolet (UV) exposure and are therefore eminently preventable [2]. Recent data suggest that especially groups with a low genetic risk benefit from UV protection [3,4] and underline the importance of aggressive prevention strategies for young target groups regarding indoor [5] and outdoor [6] UV exposure.

Unhealthy behavior in regard to UV exposure is mostly initiated in early adolescence [7], often with the idea that a tan increases attractiveness [8-10]; the problems related to malignant melanoma and skin atrophy are too far in the future to fathom. Given the substantial amount of time children and adolescents spend in schools, addressing skin cancer prevention in this setting is crucial and provides a unique opportunity to implement melanoma prevention programs [11]. In an attempt to reduce UV exposure, recent experimental studies designed for young target groups aimed at promoting sunscreen use as an end point [12-15] and others used different UV protection behaviors (including avoiding sunbeds) or behavior scores [9,16-23].

Appearance-based approaches were evaluated as superior to health-based approaches in a school-based randomized trial [12], which underlines the well-understood importance of self-perceived appearance for adolescent self-esteem [24,25].

Photoaging interventions specifically, in which a self-portrait (ie, a selfie) is altered to predict future appearance, indicated effectiveness in various behavioral change settings, including smoking cessation [26-28] or prevention [29-33], weight loss [34], and in recent years also in UV protection interventions [35-38].

We recently introduced the free mobile app Sunface, which takes advantage of the broad availability of smartphones and adolescents' interest in their appearance by photoaging and 3-dimensional (3D) animation of the users' selfie based on Fitzpatrick skin type and individual UV protection behavior [39]. Afterward, the app explains the visual results, provides guideline recommendations on sun protection plus the ABCDE rule for melanoma self-detection (assess border irregularity, color variety, diameter, and evolution [40]), and offers sharing options via photo or video (Multimedia Appendix 1).

To integrate this photoaging app for melanoma prevention in the school-based setting and to investigate how it would be perceived by adolescents who are most amenable to appearance-based interventions [24], we tested its effectiveness by the use of the mirroring approach in a pilot study. Mirroring means that the students' altered 3D selfies on mobile phones or tablets were "mirrored" via a projector in front of the whole grade. This approach was previously introduced by our group in the tobacco prevention setting [30].

Methods

Participants

We included a total sample of 205 German secondary school students of both sexes in the age group of 13-19 years in our cross-sectional pilot study (median 15 years, SD 1.36; 111/205, 54.1% male; 93/205, 45.9% female) attending the two most common school types in Germany (grammar school: n=136, 66.4%; general comprehensive school: n=69, 33.6%). Almost all participants (201/205, 98.5%) owned a smartphone.

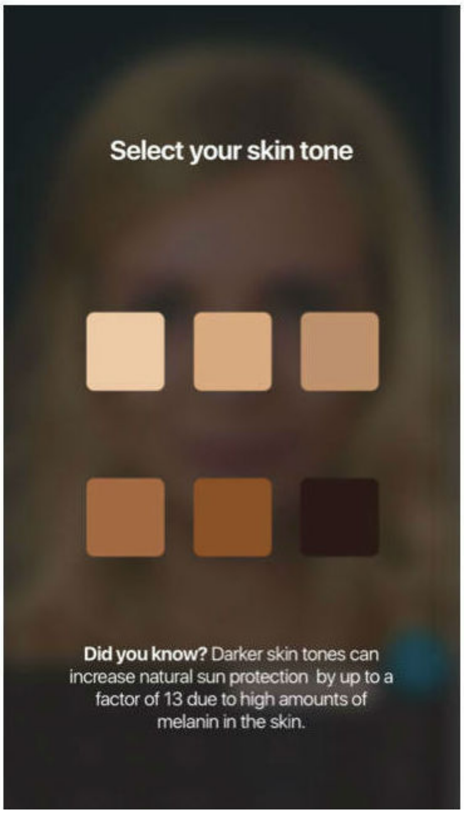
From a risk profile standpoint, 50.7% (104/205) of the participants had a Fitzpatrick skin type of 1 or 2 [41]; current sunbed use was reported by 5.9% (12/205) [42], and 77.0% (157/205) remembered having at least one sunburn in the past [43].

Intervention

The mirroring approach was implemented by medical students from the Education Against Tobacco nonprofit organization attending the University of Essen in Germany [31,44]. To increase familiarity with the photoaging app (called Sunface) and students' participation in the mirroring intervention, students were asked to download the app before our visit, via a letter 1 week in advance. By this means, 36.3% (74/205) had the app on their mobile phones when we visited the schools.

In the first 15-minute phase, we used the displayed face of one student volunteer to show the app's altering features to the peer group, providing an incentive for the rest of the class to test the app. In front of their peers and teachers, students could interact with their own animated face via touch (coughing, sneezing, etc) and display their future self based on their skin type (Figure 1) or use of sun protection (Figure 2) or tanning beds (Figure 3) at 5, 10, 15, 20, or 25 years in the future. Multiple device displays can be projected simultaneously, which we used to consolidate the altering measures with graphics (eg, to explain skin atrophy and solar elastosis). We implemented mirroring with a Galaxy Tab A tablet computer (Samsung, Seoul, South Korea) via Apple's AirPlay interface (Apple Inc) using the app Mirroring360 (Splashtop Inc) for the Android operating system (Google Inc).

Figure 1. Start of the app: the user picks their Fitzpatrick skin type.



Select your skin tone

Did you know? Darker skin tones can increase natural sun protection by up to a factor of 13 due to high amounts of melanin in the skin.

Figure 2. Effect view of the app: 25 years of aging with applied sun protection.

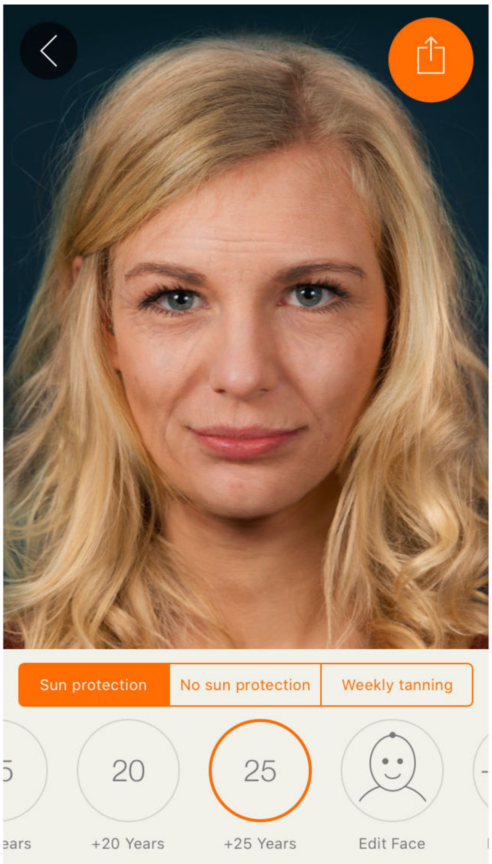


Figure 3. Effect view of the app: weekly tanning for 25 years (maximum effect) with a total of 3 actinic keratoses visible, multiple solar lentigines, age spots, and prominent solar elastosis.

In the second 15-minute phase, we encouraged students to try the app on their own device or one of the tablet computers provided for students who did not own a smartphone or did not download the app. Both the provided tablet computers and the students' own smartphones were connected to the projector.

Data Collection

We gathered the students' sociodemographic data (sex, age, school type) and their risk profile (skin type, sex, age, sunburn in the past, sunbed use) directly after the intervention via an anonymous survey. We captured their reactions toward the intervention via 6 items on a 5-point Likert scale (ranging from 1=strongly agree to 5=strongly disagree): (1) change of intentions (2 items: indoor vs outdoor tanning); and (2) perceived reactions of the peer group on change in attractiveness (2 items: indoor vs outdoor tanning), whether they perceived the intervention as fun (1 item), and perceived effects of the app as realistic (1 item). The items have been previously used in other published studies [30,45] and were pretested in advance in accordance with the guidelines for good epidemiologic practice [46].

Ethics Approval

The study received ethics approval from the ethics committee at the University of Essen (17-7587-BO).

Results

We analyzed all data for the group as a whole (Figure 4) but also to learn about how well the app was received by different Fitzpatrick skin types (Figure 5), sex (Figure 6), and age groups (Figure 7). All figures are based on the data provided in Multimedia Appendix 2 and give total numbers rather than percentages to adjust for the 5-point Likert scale design of the items.

Figure 4. Overall results of the whole sample. 3D: 3-dimensional.

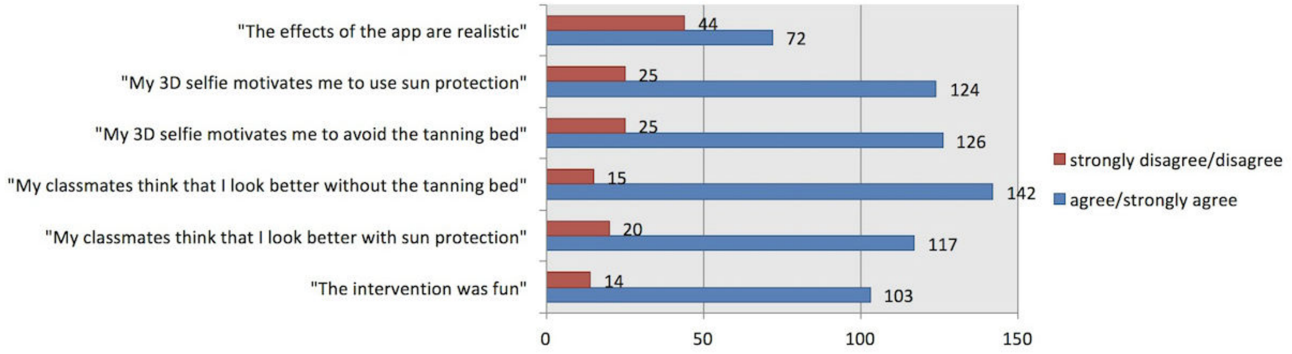


Figure 5. Results by Fitzpatrick skin types (st): 1-2 vs 3-6. 3D: 3-dimensional.

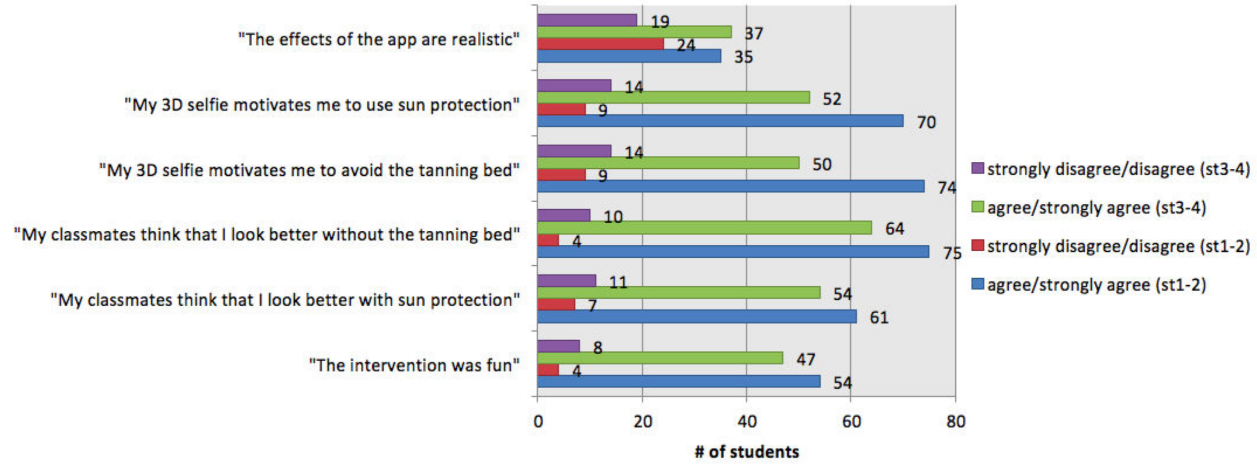


Figure 6. Results by sex: females vs males. 3D: 3-dimensional.

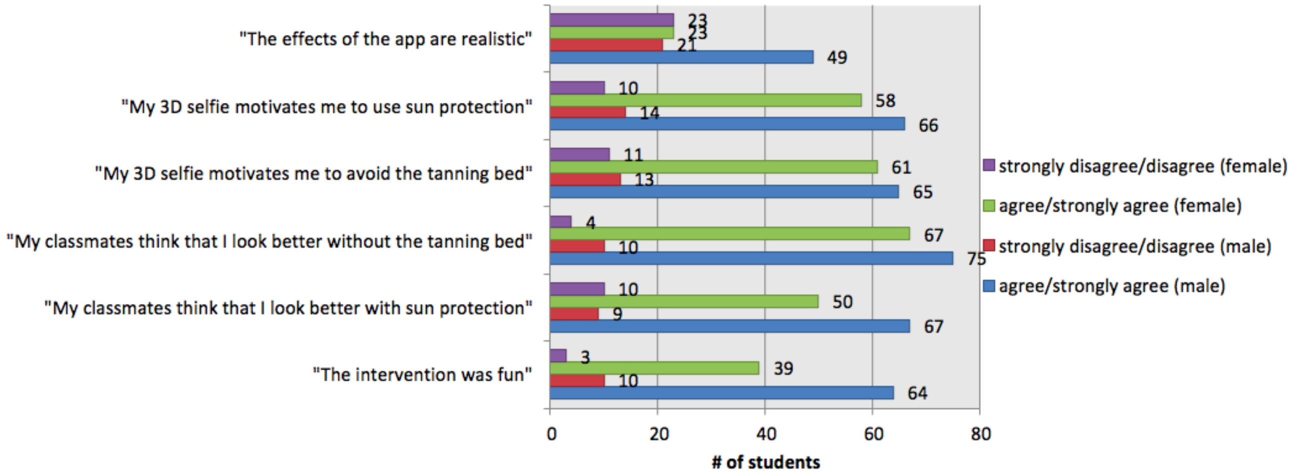
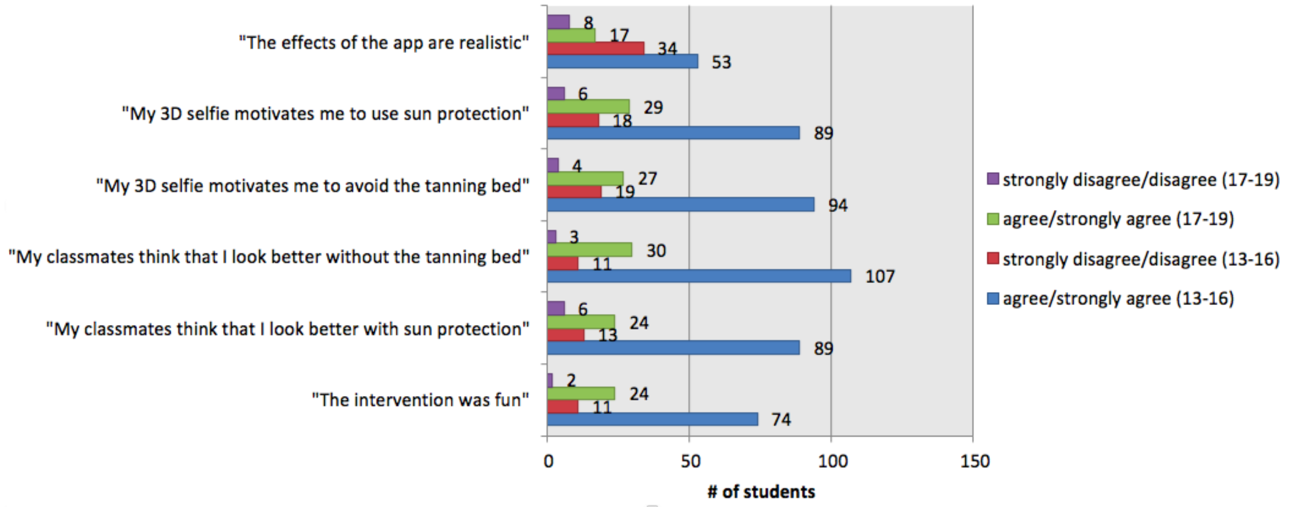


Figure 7. Results by age group: 13-16 years vs 17-19 years. 3D: 3-dimensional.



Realism of the Created Selfies

In our sample, we measured overall agreement with the subjective realism of the created selfies (72/202, 35.6% strongly agreed or agreed on realism, while n=44, 21.8% disagreed or strongly disagreed, and 42.6% (n=86) were not sure; Figure 4). These results varied notably in males (n=49, 44.5% agreement; n=21, 19.1% disagreement) versus females (n=23, 25.3% agreement; n=23, 25.3% disagreement; Figure 6) but not in other subgroups.

Motivation to Reduce Ultraviolet Exposure

We measured more than 60% agreement in both items that measured motivation to reduce UV exposure and only 12.5% disagreement (n=126, 63.0% agreed or strongly agreed that their 3D selfie motivated them to avoid the tanning bed and n=124, 61.7% to increase use of sun protection). Only 25 (12.5%) disagreed or strongly disagreed with this item and with increased use of sun protection in our sample. The perceived effect on motivation was larger in participants with Fitzpatrick skin types 1-2 in both tanning bed avoidance (n=74, 71.8% agreement in skin types 1-2 vs n=50, 53.8% agreement in skin types 3-6) and increased use of sun protection (n=70, 68.0% agreement in skin types 1-2 vs n=52, 55.3% agreement in skin types 3-6; Figure 5), and also positively correlated with higher age (Figure 7, Multimedia Appendix 2).

Perceived Subjective Norm During the Mirroring Intervention

The 2 items measuring the reactions of the peer group toward the individual selfie showed positive peer pressure in regard to both use of sun protection (n=117, 57.9%) and tanning bed avoidance (n=142, 72.1%; Figure 4). The subjective norm on decreasing UV exposure in order to look more attractive was perceived by a higher percentage of participants with Fitzpatrick skin types 1-2 than types 3-6, especially for tanning beds (Figure 5, Multimedia Appendix 2).

Global Feedback

A majority of participants claimed that they perceived the intervention as fun (n=103, 51.2% agreement vs n=14, 7.0% disagreement), and the fraction agreeing that the intervention was fun was at least two times larger than the fraction disagreeing throughout all subgroups. However, this perception was notably more prevalent in male participants (n=64, 59.3%) than in female participants (n=39, 42.4%; Figure 6). A total of 79 (39.1%) of participants reported that they would try the app again later on and 57 (28.4%) planned to show the app to another person after school (Multimedia Appendix 2).

Discussion

While multiple planned and completed trials on skin cancer prevention apps have been indexed in PubMed [47-51], to our knowledge this is the first implementation of an app-based intervention to prevent melanoma in the school setting.

Our data suggest that such an intervention is effective in changing the predictors of behavior in young risk groups and introduces a way of yielding peer-group effects in accordance with the theory of planned behavior.

Interpretation

Available data on appearance-based behavioral change settings reveal that photoaging interventions appear to be more effective for girls [31]. In our sample, the female participants perceived the intervention as less realistic in comparison with males (n=23, 25.3% agreement vs n=23, 25.3% disagreement for females) versus males (n=49, 44.5% agreement vs n=21, 19.1% disagreement; Figure 6). However, it is notable that the relatively low perception of realism by both sexes was not accompanied by low scores in the global feedback, motivation to change behavior, or subjective norm categories. Even though the questionnaires were anonymous, which reduces confounding effects in surveys, we cannot entirely rule out that the self-reported data may have been influenced by a social desirability bias.

Previous publications stressed the relevance of targeting especially fair-skinned individuals with Fitzpatrick skin type of 1 or 2, as these are less protected against UV radiation and thus have a significantly greater risk for skin cancer [1]. In our sample, the perceived effect on motivation was larger in participants with Fitzpatrick skin types 1-2 in both tanning bed avoidance (n=74, 71.8% agreement in skin types 1-2 vs n=50, 53.8% agreement in skin types 3-6) and increased use of sun protection (n=70, 68.0% agreement in skin types 1-2 vs n=52, 55.3% agreement in skin types 3-6; Figure 5, Multimedia Appendix 2), which could be explained by the Sunface app having stronger altering effects in lighter skin than in darker skin types.

Our data also indicate that photoaging interventions have a higher impact in late adolescence (17-19 years) than in early adolescence (13-16 years), but both age groups appear to benefit (Multimedia Appendix 2). We hypothesize that this effect is due to the higher relevance of wrinkle formation (ie, due to solar elastosis) in late adolescence than in early adolescence.

Conclusions

We present a novel way of integrating photoaging in school-based melanoma prevention, which affects the students' peer group, considers the predictors of UV exposure in accordance with the theory of planned behavior, and is particularly effective in changing behavioral predictors in fair-skinned adolescents (Fitzpatrick skin types 1 and 2). Further research is necessary to evaluate the intervention's prospective effects on adolescents of various cultural backgrounds.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Shared video of the Sunface app with 15 years of no sun protection shown on a 3D-animated selfie.

[MP4 File (MP4 Video), 1MB - jmir_v19i9e319_app1.mp4]

Multimedia Appendix 2

Questionnaire results.

[PDF File (Adobe PDF File), 18KB - jmir_v19i9e319_app2.pdf]

References

1. Whiteman DC, Green AC, Olsen CM. The growing burden of invasive melanoma: projections of incidence rates and numbers of new cases in six susceptible populations through 2031. J Invest Dermatol 2016 Jun;136(6):1161-1171 [FREE Full text] [doi: 10.1016/j.jid.2016.01.035] [Medline: 26902923]
2. The Lancet Editorial Board. Skin cancer: prevention is better than cure. Lancet 2014 Aug 09;384(9942):470. [doi: 10.1016/S0140-6736(14)61320-9] [Medline: 25110266]
3. Janda M, Soyer P. Greater precision in melanoma prevention. JAMA Dermatol 2017 Jan 01;153(1):18-19. [doi: 10.1001/jamadermatol.2016.3472] [Medline: 27829087]
4. Watts CG, Madronio C, Morton RL, Goumas C, Armstrong BK, Curtin A, et al. Clinical features associated with individuals at higher risk of melanoma: a population-based study. JAMA Dermatol 2017 Jan 01;153(1):23-29. [doi: 10.1001/jamadermatol.2016.3327] [Medline: 27829101]
5. Guy GP, Zhang Y, Ekwueme DU, Rim SH, Watson M. The potential impact of reducing indoor tanning on melanoma prevention and treatment costs in the United States: an economic analysis. J Am Acad Dermatol 2017 Feb;76(2):226-233. [doi: 10.1016/j.jaad.2016.09.029] [Medline: 27939556]
6. Karimkhani C, Green AC, Nijsten T, Weinstock MA, Dellavalle RP, Naghavi M, et al. The global burden of melanoma: results from the Global Burden of Disease Study 2015. Br J Dermatol 2017 Mar 30;177(1):134-140. [doi: 10.1111/bjd.15510] [Medline: 28369739]
7. Görig T, Diehl K, Greinert R, Breitbart EW, Schneider S. Prevalence of sun-protective behaviour and intentional sun tanning in German adolescents and adults: results of a nationwide telephone survey. J Eur Acad Dermatol Venereol 2017 Jun 02;31(6):913-1077. [doi: 10.1111/jdv.14376] [Medline: 28573745]
8. Schneider S, Diehl K, Bock C, Schlüter M, Breitbart EW, Volkmer B, et al. Sunbed use, user characteristics, and motivations for tanning: results from the German population-based SUN-Study 2012. JAMA Dermatol 2013 Jan;149(1):43-49. [doi: 10.1001/2013.jamadermatol.562] [Medline: 23069870]
9. Hillhouse J, Turrisi R, Scaglione NM, Cleveland MJ, Baker K, Florence LC. A web-based intervention to reduce indoor tanning motivations in adolescents: a randomized controlled trial. Prev Sci 2017 Feb;18(2):131-140. [doi: 10.1007/s11121-016-0698-4] [Medline: 27549602]
10. Hillhouse J, Turrisi R, Cleveland MJ, Scaglione NM, Baker K, Florence LC. Theory-driven longitudinal study exploring indoor tanning initiation in teens using a person-centered approach. Ann Behav Med 2016 Feb;50(1):48-57. [doi: 10.1007/s12160-015-9731-2] [Medline: 26370893]
11. Guy GP, Holman DM, Watson M. The important role of schools in the prevention of skin cancer. JAMA Dermatol 2016 Oct 01;152(10):1083-1084. [doi: 10.1001/jamadermatol.2016.3453] [Medline: 27580085]

12. Tuong W, Armstrong AW. Effect of appearance-based education compared with health-based education on sunscreen use and knowledge: a randomized controlled trial. J Am Acad Dermatol 2014 Apr;70(4):665-669. [doi: 10.1016/j.jaad.2013.12.007] [Medline: 24508292]
13. Craciun C, Schüz N, Lippke S, Schwarzer R. Facilitating sunscreen use in women by a theory-based online intervention: a randomized controlled trial. J Health Psychol 2012 Mar;17(2):207-216. [doi: 10.1177/1359105311414955] [Medline: 21752862]
14. Hirst NG, Gordon LG, Scuffham PA, Green AC. Lifetime cost-effectiveness of skin cancer prevention through promotion of daily sunscreen use. Value Health 2012;15(2):261-268 [FREE Full text] [doi: 10.1016/j.jval.2011.10.009] [Medline: 22433757]
15. Sakoufaki M, Stergiopoulou A, Stratigos A. Design and implementation of a health promotion program to prevent the harmful effects of ultraviolet radiation at primary school students of rural areas of Greece. Int J Res Dermatol 2017 Aug 24;3(3):306. [doi: 10.18203/issn.2455-4529.IntJResDermatol20172505]
16. Miller KA, Langholz BM, Ly T, Harris SC, Richardson JL, Peng DH, et al. SunSmart: evaluation of a pilot school-based sun protection intervention in Hispanic early adolescents. Health Educ Res 2015 Jun;30(3):371-379 [FREE Full text] [doi: 10.1093/her/cyv011] [Medline: 25801103]
17. Aarestrup C, Bonnesen CT, Thygesen LC, Krarup AF, Waagstein AB, Jensen PD, et al. The effect of a school-based intervention on sunbed use in Danish pupils at continuation schools: a cluster-randomized controlled trial. J Adolesc Health 2014 Feb;54(2):214-220. [doi: 10.1016/j.jadohealth.2013.08.011] [Medline: 24119418]
18. Holman DM, Fox KA, Glenn JD, Guy GP, Watson M, Baker K, et al. Strategies to reduce indoor tanning: current research gaps and future opportunities for prevention. Am J Prev Med 2013 Jun;44(6):672-681 [FREE Full text] [doi: 10.1016/j.amepre.2013.02.014] [Medline: 23683986]
19. Stapleton JL, Hillhouse J, Levonyan-Radloff K, Manne SL. Review of interventions to reduce ultraviolet tanning: need for treatments targeting excessive tanning, an emerging addictive behavior. Psychol Addict Behav 2017 Jun 22. [doi: 10.1037/adb0000289] [Medline: 28639816]
20. Olson AL, Gaffney C, Starr P, Gibson JJ, Cole BF, Dietrich AJ. SunSafe in the middle school years: a community-wide intervention to change early-adolescent sun protection. Pediatrics 2007 Jan;119(1):e247-e256. [doi: 10.1542/peds.2006-1579] [Medline: 17200250]
21. Miller KA, Huh J, Unger JB, Richardson JL, Allen MW, Peng DH, et al. Patterns of sun protective behaviors among Hispanic children in a skin cancer prevention intervention. Prev Med 2015 Dec;81:303-308 [FREE Full text] [doi: 10.1016/j.ypmed.2015.09.027] [Medline: 26436682]
22. Turner D, Harrison SL, Buettner P, Nowak M. Does being a “SunSmart School” influence hat-wearing compliance? An ecological study of hat-wearing rates at Australian primary schools in a region of high sun exposure. Prev Med 2014 Mar;60:107-114. [doi: 10.1016/j.ypmed.2013.12.020] [Medline: 24382297]
23. Buller DB, Andersen PA, Walkosz BJ, Scott MD, Beck L, Cutter GR. Rationale, design, samples, and baseline sun protection in a randomized trial on a skin cancer prevention intervention in resort environments. Contemp Clin Trials 2016 Jan;46:67-76 [FREE Full text] [doi: 10.1016/j.cct.2015.11.015] [Medline: 26593781]
24. Baudson TG, Weber KE, Freund PA. More than only skin deep: appearance self-concept predicts most of secondary school students’ self-esteem. Front Psychol 2016;7:1568 [FREE Full text] [doi: 10.3389/fpsyg.2016.01568] [Medline: 27803681]
25. Brinker TJ, Petri MP, Sondermann W, Gatzka M, Nakamura Y, Seeger W, et al. A dermatologists' ammunition in the war against smoking. J Med Internet Res 2017 Sep:1-3 (forthcoming). [doi: 10.2196/jmir.8743]
26. Burford O, Jiwa M, Carter O, Parsons R, Hendrie D. Internet-based photoaging within Australian pharmacies to promote smoking cessation: randomized controlled trial. J Med Internet Res 2013 Mar 26;15(3):e64 [FREE Full text] [doi: 10.2196/jmir.2337] [Medline: 23531984]
27. Burford O, Kindarji S, Parsons R, Falcoff H. Using visual demonstrations in young adults to promote smoking cessation: preliminary findings from a French pilot study. Res Social Adm Pharm 2017 May 04. [doi: 10.1016/j.sapharm.2017.04.050] [Medline: 28495124]
28. Flett K, Grogan S, Clark-Carter D, Gough B, Conner M. Male smokers' experiences of an appearance-focused facial-ageing intervention. J Health Psychol 2017 Mar;22(4):422-433. [doi: 10.1177/1359105315603477] [Medline: 26338489]
29. Brinker TJ, Seeger W. Photoaging mobile apps: a novel opportunity for smoking cessation? J Med Internet Res 2015 Jul 27;17(7):e186 [FREE Full text] [doi: 10.2196/jmir.4792] [Medline: 26215210]
30. Brinker TJ, Seeger W, Buslaff F. Photoaging mobile apps in school-based tobacco prevention: the mirroring approach. J Med Internet Res 2016 Jun 28;18(6):e183 [FREE Full text] [doi: 10.2196/jmir.6016] [Medline: 27352819]
31. Brinker TJ, Owczarek AD, Seeger W, Groneberg DA, Brieske CM, Jansen P, et al. A medical student-delivered smoking prevention program, education against tobacco, for secondary schools in Germany: randomized controlled trial. J Med Internet Res 2017 Jun 06;19(6):e199 [FREE Full text] [doi: 10.2196/jmir.7906] [Medline: 28588007]
32. Xavier LEDF, Bernardes-Souza B, Lisboa OC, Seeger W, Groneberg DA, Tran T, et al. A medical student-delivered smoking prevention program, education against tobacco, for secondary schools in Brazil: study protocol for a randomized trial. JMIR Res Protoc 2017 Jan 30;6(1):e16 [FREE Full text] [doi: 10.2196/resprot.7134] [Medline: 28137703]

33. Brinker TJ, Holzapfel J, Baudson TG, Sies K, Jakob L, Baumert HM, et al. Photoaging smartphone app promoting poster campaign to reduce smoking prevalence in secondary schools: the Smokerface Randomized Trial: design and baseline characteristics. BMJ Open 2016 Nov 07;6(11):e014288 [FREE Full text] [doi: 10.1136/bmjopen-2016-014288] [Medline: 27821601]
34. Jiwa M, Burford O, Parsons R. Preliminary findings of how visual demonstrations of changes to physical appearance may enhance weight loss attempts. Eur J Public Health 2015 Apr;25(2):283-285. [doi: 10.1093/eurpub/cku249] [Medline: 25644139]
35. Williams AL, Grogan S, Clark-Carter D, Buckley E. Impact of a facial-ageing intervention versus a health literature intervention on women's sun protection attitudes and behavioural intentions. Psychol Health 2013;28(9):993-1008. [doi: 10.1080/08870446.2013.777965] [Medline: 23527527]
36. Eastabrook S, Chang P, Taylor MF. Melanoma risk: adolescent females' perspectives on skin protection pre/post-viewing a ultraviolet photoaged photograph of their own facial sun damage. Glob Health Promot 2016 Jun 22. [doi: 10.1177/1757975916639871] [Medline: 27333916]
37. Mahler HIM, Kulik JA, Gerrard M, Gibbons FX. Effects of photoaging information and UV photo on sun protection intentions and behaviours: a cross-regional comparison. Psychol Health 2013;28(9):1009-1031. [doi: 10.1080/08870446.2013.777966] [Medline: 23537173]
38. Lo PL, Chang P, Taylor MF. Young Australian adults' reactions to viewing personalised UV photoaged photographs. Australas Med J 2014;7(11):454-461 [FREE Full text] [doi: 10.4066/AMJ.2014.2253] [Medline: 25550717]
39. Brinker TJ, Schadendorf D, Klode J, Cosgarea I, Rösch A, Jansen P, et al. Photoaging mobile apps as a novel opportunity for melanoma prevention: pilot study. JMIR Mhealth Uhealth 2017 Jul 26;5(7):e101 [FREE Full text] [doi: 10.2196/mhealth.8231] [Medline: 28747297]
40. Robinson JK, Wayne JD, Martini MC, Hultgren BA, Mallett KA, Turrisi R. Early detection of new melanomas by patients with melanoma and their partners using a structured skin self-examination skills training intervention: a randomized clinical trial. JAMA Dermatol 2016 Sep 01;152(9):979-985 [FREE Full text] [doi: 10.1001/jamadermatol.2016.1985] [Medline: 27367303]
41. Roider EM, Fisher DE. Red hair, light skin, and UV-independent risk for melanoma development in humans. JAMA Dermatol 2016 Jul 01;152(7):751-753 [FREE Full text] [doi: 10.1001/jamadermatol.2016.0524] [Medline: 27050924]
42. Ghiasvand R, Rueegg CS, Weiderpass E, Green AC, Lund E, Veierød MB. Indoor tanning and melanoma risk: long-term evidence from a prospective population-based cohort study. Am J Epidemiol 2017 Dec 01;185(3):147-156. [doi: 10.1093/aje/kww148] [Medline: 28077359]
43. Wu S, Cho E, Li W, Weinstock MA, Han J, Qureshi AA. History of severe sunburn and risk of skin cancer among women and men in 2 prospective cohort studies. Am J Epidemiol 2016 May 01;183(9):824-833 [FREE Full text] [doi: 10.1093/aje/kwv282] [Medline: 27045074]
44. Brinker TJ, Stamm-Balderjahn S, Seeger W, Klingelhöfer D, Groneberg DA. Education Against Tobacco (EAT): a quasi-experimental prospective evaluation of a multinational medical-student-delivered smoking prevention programme for secondary schools in Germany. BMJ Open 2015 Sep 18;5(9):e008093 [FREE Full text] [doi: 10.1136/bmjopen-2015-008093] [Medline: 26384722]
45. Detert H, Hedlund S, Anderson CD, Rodvall Y, Festin K, Whiteman DC, et al. Validation of sun exposure and protection index (SEPI) for estimation of sun habits. Cancer Epidemiol 2015 Dec;39(6):986-993 [FREE Full text] [doi: 10.1016/j.canep.2015.10.022] [Medline: 26547793]
46. Hoffmann W, Latza U, Terschüren C. [Guidelines and recommendations for ensuring Good Epidemiological Practice (GEP) -- revised version after evaluation]. Gesundheitswesen 2015 Mar;67(3):217-225. [doi: 10.1055/s-2004-813850] [Medline: 15789285]
47. Rodrigues AM, Sniehotta FF, Birch-Machin MA, Olivier P, Araújo-Soares V. Systematic and iterative development of a smartphone app to promote sun-protection among holidaymakers: design of a prototype and results of usability and acceptability testing. JMIR Res Protoc 2017 Jun 12;6(6):e112 [FREE Full text] [doi: 10.2196/resprot.7172] [Medline: 28606892]
48. Dorairaj JJ, Healy GM, McInerney A, Hussey AJ. Validation of a melanoma risk assessment smartphone application. Dermatol Surg 2017 Feb;43(2):299-302. [doi: 10.1097/DSS.0000000000000916] [Medline: 28165352]
49. Bert F, Gualano MR, Clemente S, Villa G, Siliquini R. [Technology and prevention in the era of mobile health: applications for cancer screening programs]. Ig Sanita Pubbl 2017;73(1):13-27. [Medline: 28428642]
50. Mar VJ, Scolyer RA, Long GV. Computer-assisted diagnosis for skin cancer: have we been outsmarted? Lancet 2017 May 20;389(10083):1962-1964. [doi: 10.1016/S0140-6736(17)31285-0] [Medline: 28534744]
51. Ngoo A, Finnane A, McMeniman E, Tan J, Janda M, Soyer HP. Efficacy of smartphone applications in high-risk pigmented lesions. Australas J Dermatol 2017 Feb 27. [doi: 10.1111/ajd.12599] [Medline: 28240347]

Abbreviations

3D: 3-dimensional

UV: ultraviolet

Edited by G Eysenbach; submitted 07.08.17; peer-reviewed by O Burford, J Makin; comments to author 24.08.17; revised version received 30.08.17; accepted 30.08.17; published 08.09.17.

Please cite as:
Brinker TJ, Brieske CM, Schaefer CM, Buslaff F, Gatzka M, Petri MP, Sondermann W, Schadendorf D, Stoffels I, Klode J
Photoaging Mobile Apps in School-Based Melanoma Prevention: Pilot Study
J Med Internet Res 2017;19(9):e319
URL: http://www.jmir.org/2017/9/e319/
doi:10.2196/jmir.8661
PMID:28887295

Original Paper

Prescribing of Electronic Activity Monitors in Cardiometabolic Diseases: Qualitative Interview-Based Study

Alice Bellicha[1,2], MSc; Sandrine Macé[3], PhD; Jean-Michel Oppert[1,4], MD, PhD

[1]Institute of Cardiometabolism and Nutrition, University Pierre et Marie Curie-Paris 6, Paris, France
[2]Laboratory of Bioengineering, Tissues and Neuroplasticity, University Paris-Est, Créteil, France
[3]Department of Marketing, Ecole Supérieure de Commerce de Paris-Europe, Paris, France
[4]Assistance Publique-Hôpitaux de Paris, Department of Nutrition, Pitie-Salpetriere University Hospital, Paris, France

Corresponding Author:
Alice Bellicha, MSc
Laboratory of Bioengineering, Tissues and Neuroplasticity
University Paris-Est
8 rue Jean Sarrail
Créteil, 94010
France
Phone: 33 42175782
Email: alice.bellicha@u-pec.fr

Abstract

Background: The prevalence of noncommunicable diseases, including those such as type 2 diabetes, obesity, dyslipidemia, and hypertension, so-called cardiometabolic diseases, is high and is increasing worldwide. Strong evidence supports the role of physical activity in management of these diseases. There is general consensus that mHealth technology, including electronic activity monitors, can potentially increase physical activity in patients, but their use in clinical settings remains limited. Practitioners' requirements when prescribing electronic activity monitors have been poorly described.

Objective: The aims of this qualitative study were (1) to explore how specialist physicians prescribe electronic activity monitors to patients presenting with cardiometabolic conditions, and (2) to better understand their motivation for and barriers to prescribing such monitors.

Methods: We conducted qualitative semistructured interviews in March to May 2016 with 11 senior physicians from a public university hospital in France with expertise in management of cardiometabolic diseases (type 1 and type 2 diabetes, obesity, hypertension, and dyslipidemia). Interviews lasted 45 to 60 minutes and were audiotaped, transcribed verbatim, and analyzed using directed content analysis. We report our findings following the Consolidated Criteria for Reporting Qualitative Research (COREQ) checklist.

Results: Most physicians we interviewed had never prescribed electronic activity monitors, whereas they frequently prescribed blood glucose or blood pressure self-monitoring devices. Reasons for nonprescription included lack of interest in the data collected, lack of evidence for data accuracy, concern about work overload possibly resulting from automatic data transfer, and risk of patients becoming addicted to data. Physicians expected future marketing of easy-to-use monitors that will accurately measure physical activity duration and intensity and provide understandable motivating feedback.

Conclusions: Features of electronic activity monitors, although popular among the general public, do not meet the needs of physicians. In-depth understanding of physicians' expectations is a first step toward designing technologies that can be widely used in clinical settings and facilitate physical activity prescription. Physicians should have a role, along with key health care stakeholders—patients, researchers, information technology firms, the public, and private payers—in developing the most effective methods for integrating activity monitors into patient care.

(J Med Internet Res 2017;19(9):e328) doi:10.2196/jmir.8107

KEYWORDS

cardiometabolic diseases; physical activity; physicians' perspectives; prescriptions; mobile health; telemedicine; mHealth; electronic activity monitors; fitness tracker; accelerometer; smart pedometer

Introduction

Physical inactivity is recognized as a leading cause of noncommunicable diseases, including cardiometabolic conditions such as type 2 diabetes, obesity, dyslipidemia, and hypertension [1,2]. Despite the well-established benefits of physical activity for preventive care and management of these diseases, it remains underprescribed by physicians [3]. The rapid expansion of mobile technology, including electronic activity monitors (EAMs), presents an opportunity for encouraging physicians to prescribe physical activity. EAMs typically track daily movement, mainly the number of steps taken, with sensors both recording acceleration and providing feedback to the user via a monitor display or a smartphone app [4]. Two recent randomized controlled trials showed that EAMs associated with individually tailored feedback may increase physical activity in individuals with overweight and obesity or type 2 diabetes [5,6]. However, a mismatch between information technology firms that are developing new technologies and the physicians who use them has been reported [7]. In addition, knowledge about the needs of physicians when prescribing EAMs to patients is very limited. The aims of this qualitative study were therefore (1) to explore how specialist physicians prescribe EAMs to patients presenting with cardiometabolic conditions, and (2) to better understand physicians' motivations for and barriers to prescribing such monitors.

Methods

Procedures and Participants

We conducted qualitative semistructured interviews to investigate the prescribing of EAMs in patient care, following guidelines from the Consolidated Criteria for Reporting Qualitative Research (COREQ) [8]. This type of analysis is used when prior research on a subject exists but is incomplete or could benefit from further description [9].

Participants were hospital physicians with expertise in management of patients with cardiometabolic diseases (ie, type 1 and type 2 diabetes, obesity, dyslipidemia, and hypertension). Recruitment took place from March-May 2016 in one public university hospital in Paris (Assistance Publique-Hôpitaux de Paris, France). We contacted 11 physicians whom we knew by email and personally invited them to participate. All agreed to participate. We used purposive sampling to achieve a varied composition in terms of their sex, experience, status, and fields of expertise (Table 1).

Table 1. Characteristics of physicians interviewed.

Code	Sex	Age range (years)	Medical specialty	Status
P1	Male	50-59	Obesity	Hospital physician, university professor
P2	Female	30-39	Obesity, diabetes	Hospital physician
P3	Female	50-59	Cardiology, obesity	Hospital physician, private practice
P4	Male	≥60	Obesity, diabetes	Hospital physician, university professor
P5	Male	≥60	Diabetes	Hospital physician, university professor
P6	Female	50-59	Diabetes	Hospital physician
P7	Female	40-49	Diabetes	Hospital physician, university professor
P8	Male	50-59	Endocrinology, dyslipidemia	Hospital physician, university professor
P9	Male	50-59	Endocrinology, dyslipidemia	Hospital physician, university professor
P10	Male	50-59	Hypertension	Hospital physician, university professor
P11	Female	50-59	Obesity	Hospital physician, private practice

Diversity within the sample of specialist physicians was important so as to take into account different opinions and further improve understanding. Among the physicians, 4 were heads of their departments at the time of the study and 1 had headed a department in the recent past. Participants received a verbal explanation on the aim of the study prior to the interviews. The study complied with standards set by the Declaration of Helsinki, and we obtained written consent from all participants.

Interviews

We conducted all interviews at work in the physicians' offices. No one was present besides the participant and the researcher conducting the interviews. Interviews lasted 45 to 60 minutes and were carried out by the first author (AB), a PhD student in sports science who had attended a training course on qualitative research prior to the study. We developed a discussion guide, including open-ended questions, prior to the first interview. We asked physicians to explain how they handled the issue of physical activity with their patients. As such, we examined 4 stages of physical activity counselling: (1) initial assessment of physical activity, (2) prescription, (3) patient education, and (4) follow-up evaluation [10]. Next, we asked physicians to describe their experiences with EAMs in routine clinical care, motivations for and barriers to prescription, and their expectations. All interviews were audiotaped with the consent of participants and transcribed verbatim. Field notes were made during the interviews to facilitate data analysis and interpretation [8].

Data Analysis

We used thematic analysis via a directed approach [9]. After completion of the interviews, repeated reading of transcripts enabled familiarity with the data. Then, 1 investigator coded the transcripts according to predefined categories (initial evaluation, prescription, education, and follow-up evaluation). Data that could not be coded using the initial categories were identified and later analyzed to determine whether they represented a new category. We discussed the coding process with each other throughout the analysis. Physicians' comments are presented verbatim (translated from French into English) in the Results and are identified by the physicians' code (Table 1).

Results

Prescribing of Physical Activity

Interviewed physicians consistently recommended physical activity to their patients, mainly orally. Several barriers to prescribing physical activity emerged.

Lack of Evaluation Tools

Prescribing of physical activity was seen as requiring a specific diagnosis, often not available to physicians; they all reported difficulties in accurately assessing physical activity, especially its intensity.

> *Patients are not aware of their physical activity. They think they perform some physical activity just because they walk. But in reality, that's not physical activity. I want them to sweat.* [P10]

Opportunity Cost

Prescribing physical activity was described as a difficult, time-consuming task, especially when compared with prescribing a drug. Discussing physical activity with rather reluctant patients and reinforcing their motivation for physical activity would take more time and require greater involvement.

> *It's much more difficult to give advice on moving 30 minutes a day than to say "Take this pill."* [P2]

In addition, accompanying patients in their daily practice of physical activity, although of prime importance, is not a main task for physicians. They deplored the lack of available solutions for patient follow-up.

> *Apart from pedometers, we have nothing to propose. There is no physical activity instructor in my department, and this is lacking.* [P7]

Perceived Risk

Prescribing physical activity was considered more risky than prescribing drugs, mainly because physical activity cannot generally be accepted as a cure compared with drug treatment, the symbolic nature of which was mentioned.

> *You can provide a solution in less than 5 minutes with a drug prescription, just by saying "Try this and you'll get better." It's like selling dreams, a cure, in just one sentence, it's extraordinary. With physical activity, you can't do that, unless maybe you find tools that tell you how to prescribe it.* [P10]

Also, loss of credibility possibly resulting from lack of efficacy of physical activity appeared to be a barrier to prescription.

> *When we say to patients "Your blood glucose level will go down" and this does not turn out to be the case, we lose credibility. Whereas with medication, we know it's going to work.* [P8]

Prescribing of Electronic Activity Monitors

Experience With Electronic Activity Monitors

Physicians agreed that an increasing number of patients own an EAM.

> *Many patients have already downloaded an app for tracking physical activity.* [P11]

Most physicians said that they spent time analyzing data with patients.

> *I ask where they stand, if it helps, I encourage them, look at the results and comment on them.* [P7]

But only 4 physicians had already recommended either an EAM or the pedometer included in mobile phones, and only on rare occasions.

Perceived Benefits

Physicians acknowledged that EAMs might help patients to assess their physical activity level and might potentially motivate them, providing novelty and a recreational aspect in the context of long-term management of chronic disease. Some physicians also described EAMs as a means of improving the patient-doctor relationship.

> *It enables us to discuss something concrete.* [P2]

Lack of Clinical Utility

Barriers to prescription of EAMs included lack of evidence on data validity and reliability, and lack of interest in data collected. Physicians felt that data collected by EAMs are unsuitable for meeting the goals they define for the patient.

> *I can't see the benefit of an electronic activity monitor for me. The information isn't useful.* [P10]
>
> *Patients who bring their connected data, I have no time for that. Their number of steps...it's not a goal I've defined.* [P10]

Physicians expected an accurate measurement of relevant outcome (eg, physical activity duration and intensity, time spent sitting) that would be presented in the form of summary scores readily understandable by both the patient and physician. They also wished to personalize criteria such as physical activity goal, type of data synthesized, period of analysis, and prompts and feedback sent to the patient.

> *Positive feedback, okay, but not all the time.* [P4]

Some physicians also suggested that physical activity data should be translated into motivating benefits, such as long-term health improvement (life-years gained), better short-term disease management (lower blood glucose, lower insulin dose, etc), and improved well-being.

Learning and Searching Cost

Physicians reported a lack of knowledge of available devices and difficulty in keeping aware of the continuously growing newly marketed devices.

> *I can't be asked to learn about the available monitors. I don't have time. It's not the doctor's role.* [P8]

Convenience, ease of use, and prior knowledge of EAMs were strongly awaited. Physicians also feared work overload that would result from automatic data transfer, that is, data that would be automatically transferred electronically to the physician between medical consultations.

> *The problem with connected devices is the additional workload. Patients send me emails all the time.* [P2]

In contrast, they did state that they wanted to discuss the data during the medical consultation.

> *The patient must bring data, otherwise I won't have time to analyze them.* [P1]

Monitoring and Privacy Costs

Physicians pointed out the risk of patients becoming addicted to data.

> *The main problem with self-monitoring is the risk of addiction.* [P3]
>
> *The question is, how can I correctly use the electronic activity monitor and know I'm making progress, without becoming addicted to it?* [P5]

They also highlighted the risk of potential control of insurance coverage over the patients.

> *If physical activity data are going to be sent to private insurers for a bonus-malus contract, then the answer is no.* [P5]

Financial Issues

The high cost of devices was considered a barrier to their prescription. Most physicians recommended a selling price under €50 (about US $56) so as to be affordable for patients, and by analogy with the price of blood pressure self-monitoring devices. However, they explained that the monitors must have real added value to justify the expense.

> *Patients are not willing to pay high prices. The monitor must have real added value.* [P6]

Reimbursement by public or private insurance was not viewed as essential. Physicians believed that renting the EAMs for a limited time period rather than purchasing it could decrease costs and risk of addiction to data, and emphasize its educational role.

Prescription of Self-Monitoring Devices

Blood pressure or blood glucose self-monitoring devices were frequently prescribed, with the principal aim of improving diagnosis and optimizing drug treatment.

> *Self-monitoring of blood glucose provides real information on whether the insulin dose is the right one.* [P2]

When devices that are connected were available, physicians recommended them because they seem easier to use in everyday life settings. However, they were reticent about automatic data transfer. One physician explained that, in the context of chronic diseases, automatic data transfer goes against the principle of patient autonomy.

> *Transmitting blood glucose to a nurse who will tell the patient what to do, it's ridiculous, even more so in diabetes, that requires immediate action.* [P5]

Discussion

Principal Findings

Activity monitors are becoming increasingly popular among the general public. However, our results suggest that hospital physicians with expertise in management of patients with cardiometabolic diseases have not yet integrated EAMs into routine clinical care, which contrasts with their frequent prescription of other types of self-monitoring devices, such as blood pressure or blood glucose monitors.

While most physicians had never recommended EAMs, they acknowledged their potential to increase patient motivation through precise quantification of physical activity. However, they were concerned about data validity, which has been shown in several published studies to be insufficient [11]. EAMs are consumer-grade monitors often validated by the company only after market launching, without external validation [12].

The physicians we interviewed also questioned the clinical utility of recording step counts. Physicians followed current physical activity guidelines that recommend a given duration and intensity of physical activity but do not recommend a cutoff for steps taken per day [13-15]. Given the inability of patients to accurately estimate intensity, physicians expected this outcome to be measured by EAMs. It is surprising that companies have not yet designed EAMs measuring physical activity intensity, since it could be easily obtained via minute-by-minute analysis of step counts. Walking cadence (ie, number of steps/min) is recognized as a valuable reflection of intensity, and thresholds have been proposed to categorize intensity based on cadence [16]. Cadence has the advantage of being easily interpretable by patients and physicians. Real-time feedback of walking cadence would improve patients' perception of intensity, highlighting the potential educational role of EAMs [17]. Moreover, by helping physicians to assess physical activity, EAMs could overcome a major difficulty in prescribing physical activity.

Beyond EAMs' role in data collection, physicians attach great importance to the feedback provided by EAMs. They expected this to be easily understandable by both patient and physician, and to be presented in the form of summary scores over personalized periods of time. They also suggested that physical activity data be translated into short-term benefits related either to improved disease management (eg, better regulation of blood glucose, reduction in insulin dose) or to improved well-being. Emphasizing the benefits of physical activity would provide patients with immediate concrete rewards [18], which are known to have higher priority than greater but delayed rewards, and could therefore improve adherence to long-term lifestyle changes [19,20]. This proposition points out a major difference in how physicians use EAMs and blood glucose or blood pressure self-monitoring devices. The latter devices provide direct, more accurate and complete measures of health outcomes than does traditional monitoring. Their main benefits for physicians have been to optimize drug prescriptions and, for patients, to make appropriate treatment choices and motivate lifestyle changes [21]. Therefore, by analogy with widely adopted blood glucose and blood pressure monitors, EAMs could be designed to provide feedback concerning the benefits of physical activity for health outcomes. Such feedback would decrease the perceived risk associated with prescribing of physical activity. A simple translation of physical activity data into health benefits, as suggested by physicians, appears to be a feasible strategy that could be rapidly implemented by companies. Some authors have predicted that, in the near future, a single device will have the capacity to monitor a range of data, including both physical activity and relevant medical data [22], and will provide patients and physicians with a direct measure of physical activity benefits.

A striking finding of this study was that the learning and searching costs associated with use of EAMs prevented physicians from prescribing them. They deplored a lack of information about available EAMs and difficulties with choosing between the ever-growing number of devices on the market. In contrast, they appeared better informed about blood pressure or blood glucose monitors, which are regulated medical devices [23]. Lists of devices that have been independently validated for use in clinical practice are freely available [24], as are guidelines on how to use them for assessment and management of diabetes and hypertension, which is not yet the case for EAMs [21,25]. A broader adoption of EAMs will necessarily involve lowering learning and searching costs for physicians. Over the short term, the effort to gather information on the validity, features, or cost of commercially available EAMs could be assigned to other health care professionals, such as physical activity instructors. Over a longer term, EAMs that meet validity and effectiveness requirements of medical device regulations could be marketed, and the best practices to be shared between physicians and patients would be defined.

Physicians highlighted monitoring risks that patients might experience when using EAMs, especially that of addiction to data. Such risks have been described in patients using blood glucose and blood pressure self-monitoring devices [26,27]. Thus far, studies have suggested a decrease in adherence to EAMs over time, without mentioning the risk of addiction to physical activity data [22]. The physicians interviewed here supported the idea that patients should rent EAMs for a limited time period rather than purchasing them. Renting has the advantage of decreasing both the risk of addiction to data and financial costs that physicians consider too high to be affordable for patients of low socioeconomic status [28]. Physicians also warned against use of data by private insurers who might penalize insufficiently active patients. Such contracts have recently been authorized in some countries [29], although they are not authorized in others. This warning points to the critical issue of data privacy, now recognized as a priority by companies and regulators [29].

Finally, and surprisingly, all of the physicians we interviewed strongly opposed automatic data transfer, not only because of time constraints, but also because they considered that developing patient autonomy and self-care ability is a major aspect of patient education in the context of chronic diseases. The ability of wearable technology to transmit data to the physician is usually presented as attractive for clinical applications [30]. However, physicians have neither the time nor the desire to receive physical activity data, preferring to discuss data with the patient during the more traditional context of a medical consultation. Our data highlight the need for companies to work closely with physicians to determine when the contact with connected data is clinically useful, which may vary according to the disease, the type of data collected, and individual preferences [7].

Strengths and Limitations

First, this qualitative study provides a perspective on attitudes of hospital physicians regarding EAMs and their integration into patient care. To our knowledge, this is the first study that specifically targeted hospital physicians with expertise in cardiometabolic diseases, a set of conditions with increasing prevalence worldwide. Hospital physicians are opinion leaders in the medical community and influence the prescription habits of their colleagues [31]. Second, we explored barriers to physical activity prescription, a necessary first step in understanding whether and how EAMs might encourage physicians to prescribe physical activity.

Our study has some limitations. First, most physicians interviewed were aged 50 years or older. Younger physicians may have different beliefs regarding EAMs, since adoption of new technologies is higher at younger ages [32]. Second, the relatively small number of participants was also a limitation, although it is acknowledged that the number of participants can be reduced when the degree of expertise increases [33]. Third, the physicians we interviewed were working in a university setting in Paris, France, and the findings may not directly apply to other medical settings around the world. Assessing the views of other profiles of physicians about EAM use would be useful. Fourth, we analyzed data using a directed content analysis, the main limitation being that researchers approach data with an informed and potential bias [9]. To limit bias, we asked only open-ended questions, so as to allow unexpected but relevant themes to emerge. In addition, we discussed coding of transcripts with each other throughout the analysis.

Conclusions

The increased use of EAMs provides a timely opportunity to encourage prescribing of physical activity. EAMs have the potential to improve patient education and motivation through better assessment of physical activity, to enable a more precise prescription of physical activity, and to reinforce the patient-doctor relationship. However, hospital physicians with expertise in management of cardiometabolic diseases have not yet adopted EAMs. To do so, numerous barriers must be overcome. Important adaptations could be rapidly achieved (eg, measuring physical activity intensity through walking cadence, emphasizing health benefits), while others require more time and effort from key health care stakeholders (eg, defining best practices, regulating data privacy). This study pointed out questions related to the most effective use of EAMs for management of chronic diseases that should be explored in future studies.

Acknowledgments

AB, JMO, and SM designed the study and discussion guide. AB conducted the interviews, analyzed the data, synthesized results, and drafted the manuscript. SM and JMO supervised data analysis and contributed to writing the manuscript.

This research did not receive any specific grant from funding agencies in the public, commercial, or nonprofit sectors. AB is a PhD student supported by the Institute of Cardiometabolism and Nutrition (ICAN, Paris, France), and the work was part of her PhD.

We would like to acknowledge all physicians who participated in the interviews.

Conflicts of Interest

SM is the Scientific Director of the Internet of Things (IoT) Chair at ESCP Europe, sponsored by Schneider Electric, Société Générale Insurance, and Valeo. JMO reports research grants from Withings, outside the submitted work.

References

1. Lee I, Shiroma EJ, Lobelo F, Puska P, Blair SN, Katzmarzyk PT, Lancet Physical Activity Series Working Group. Effect of physical inactivity on major non-communicable diseases worldwide: an analysis of burden of disease and life expectancy. Lancet 2012 Jul 21;380(9838):219-229 [FREE Full text] [doi: 10.1016/S0140-6736(12)61031-9] [Medline: 22818936]
2. Physical activity Guidelines Advisory Committee. Physical Activity Guidelines Advisory Committee Report. Washington, DC: US Department of Health and Human Services; 2008.
3. Barnes PM, Schoenborn CA. Trends in adults receiving a recommendation for exercise or other physical activity from a physician or other health professional. NCHS Data Brief 2012 Feb(86):1-8 [FREE Full text] [Medline: 22617014]
4. Hickey AM, Freedson PS. Utility of consumer physical activity trackers as an intervention tool in cardiovascular disease prevention and treatment. Prog Cardiovasc Dis 2016;58(6):613-619. [doi: 10.1016/j.pcad.2016.02.006] [Medline: 26943981]
5. Cadmus-Bertram LA, Marcus BH, Patterson RE, Parker BA, Morey BL. Randomized trial of a Fitbit-based physical activity intervention for women. Am J Prev Med 2015 Sep;49(3):414-418. [doi: 10.1016/j.amepre.2015.01.020] [Medline: 26071863]
6. Hochberg I, Feraru G, Kozdoba M, Mannor S, Tennenholtz M, Yom-Tov E. Encouraging physical activity in patients with diabetes through automatic personalized feedback via reinforcement learning improves glycemic control. Diabetes Care 2016 Apr;39(4):e59-e60. [doi: 10.2337/dc15-2340] [Medline: 26822328]
7. Rudin RS, Bates DW, MacRae C. Accelerating innovation in health IT. N Engl J Med 2016 Sep 01;375(9):815-817. [doi: 10.1056/NEJMp1606884] [Medline: 27579633]
8. 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 [FREE Full text] [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
9. Hsieh H, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res 2005 Nov;15(9):1277-1288. [doi: 10.1177/1049732305276687] [Medline: 16204405]
10. Garber CE, Blissmer B, Deschenes MR, Franklin BA, Lamonte MJ, Lee I, American College of Sports Medicine. American College of Sports Medicine position stand. Quantity and quality of exercise for developing and maintaining cardiorespiratory, musculoskeletal, and neuromotor fitness in apparently healthy adults: guidance for prescribing exercise. Med Sci Sports Exerc 2011 Jul;43(7):1334-1359. [doi: 10.1249/MSS.0b013e318213fefb] [Medline: 21694556]
11. Evenson KR, Goto MM, Furberg RD. Systematic review of the validity and reliability of consumer-wearable activity trackers. Int J Behav Nutr Phys Act 2015;12(1):159 [FREE Full text] [doi: 10.1186/s12966-015-0314-1] [Medline: 26684758]
12. Sperlich B, Holmberg H. Wearable, yes, but able…?: it is time for evidence-based marketing claims!. Br J Sports Med 2016 Dec 16:1240 [FREE Full text] [doi: 10.1136/bjsports-2016-097295] [Medline: 27986762]
13. Thompson D, Batterham AM, Peacock OJ, Western MJ, Booso R. Feedback from physical activity monitors is not compatible with current recommendations: a recalibration study. Prev Med 2016 Oct;91:389-394 [FREE Full text] [doi: 10.1016/j.ypmed.2016.06.017] [Medline: 27330025]
14. US Department of Health and Human Services. Physical Activity Guidelines for Americans. Washington, DC: US Department of Health and Human Services; 2008.
15. O'Donovan G, Blazevich AJ, Boreham C, Cooper AR, Crank H, Ekelund U, et al. The ABC of Physical Activity for Health: a consensus statement from the British Association of Sport and Exercise Sciences. J Sports Sci 2010 Apr;28(6):573-591. [doi: 10.1080/02640411003671212] [Medline: 20401789]

16. Tudor-Locke C, Rowe DA. Using cadence to study free-living ambulatory behaviour. Sports Med 2012 May 01;42(5):381-398. [doi: 10.2165/11599170-000000000-00000] [Medline: 22462794]
17. Bouchard DR, Langlois M, Boisvert-Vigneault K, Farand P, Paulin M, Baillargeon J. Pilot study: can older inactive adults learn how to reach the required intensity of physical activity guideline? Clin Interv Aging 2013;8:501-508 [FREE Full text] [doi: 10.2147/CIA.S42224] [Medline: 23662053]
18. Reach G. Can technology improve adherence to long-term therapies? J Diabetes Sci Technol 2009 May 01;3(3):492-499 [FREE Full text] [doi: 10.1177/193229680900300313] [Medline: 20144287]
19. Kahneman D. Thinking Fast and Slow. Wallingford, UK: Penguin; 2012.
20. Epstein LH, Salvy SJ, Carr KA, Dearing KK, Bickel WK. Food reinforcement, delay discounting and obesity. Physiol Behav 2010 Jul 14;100(5):438-445. [doi: 10.1016/j.physbeh.2010.04.029] [Medline: 20435052]
21. American Diabetes Association. 5. Glycemic targets. Diabetes Care 2016 Jan;39(Suppl 1):S39-S46. [doi: 10.2337/dc16-S008] [Medline: 26696679]
22. Piwek L, Ellis DA, Andrews S, Joinson A. The rise of consumer health wearables: promises and barriers. PLoS Med 2016 Feb;13(2):e1001953 [FREE Full text] [doi: 10.1371/journal.pmed.1001953] [Medline: 26836780]
23. Istepanian RSH. Mobile applications for diabetes management: efficacy issues and regulatory challenges. Lancet Diabetes Endocrinol 2015 Dec;3(12):921-923. [doi: 10.1016/S2213-8587(15)00414-3] [Medline: 26545304]
24. dabl Educational Trust. Blood pressure monitors - validations, papers and reviews. Manchester, UK: dabl Educational Trust Ltd; 2015. URL: http://www.dableducational.org/ [accessed 2017-05-24] [WebCite Cache ID 6qhQDkf5U]
25. Parati G, Stergiou GS, Asmar R, Bilo G, de Leeuw P, Imai Y, et al. European Society of Hypertension guidelines for blood pressure monitoring at home: a summary report of the Second International Consensus Conference on Home Blood Pressure Monitoring. J Hypertens 2008 Aug;26(8):1505-1526. [doi: 10.1097/HJH.0b013e328308da66] [Medline: 18622223]
26. Fletcher BR, Hinton L, Hartmann-Boyce J, Roberts NW, Bobrovitz N, McManus RJ. Self-monitoring blood pressure in hypertension, patient and provider perspectives: a systematic review and thematic synthesis. Patient Educ Couns 2016 Feb;99(2):210-219. [doi: 10.1016/j.pec.2015.08.026] [Medline: 26341941]
27. Peel E, Parry O, Douglas M, Lawton J. Blood glucose self-monitoring in non-insulin-treated type 2 diabetes: a qualitative study of patients' perspectives. Br J Gen Pract 2004 Mar;54(500):183-188 [FREE Full text] [Medline: 15006123]
28. Sommer I, Griebler U, Mahlknecht P, Thaler K, Bouskill K, Gartlehner G, et al. Socioeconomic inequalities in non-communicable diseases and their risk factors: an overview of systematic reviews. BMC Public Health 2015 Sep 18;15:914 [FREE Full text] [doi: 10.1186/s12889-015-2227-y] [Medline: 26385563]
29. Armstrong S. What happens to data gathered by health and wellness apps? BMJ 2016 Jun 23;353:i3406. [doi: 10.1136/bmj.i3406] [Medline: 27338775]
30. Murray E. eHealth: where next? Br J Gen Pract 2014 Jul;64(624):325-326 [FREE Full text] [doi: 10.3399/bjgp14X680365] [Medline: 24982470]
31. Sbarbaro JA. Can we influence prescribing patterns? Clin Infect Dis 2001 Sep 15;33(Suppl 3):S240-S244. [doi: 10.1086/321856] [Medline: 11524726]
32. Jamoom E, Beatty P, Bercovitz A, Woodwell D, Palso K, Rechtsteiner E. Physician adoption of electronic health record systems: United States, 2011. NCHS Data Brief 2012 Jul(98):1-8 [FREE Full text] [Medline: 23050588]
33. Mason M. Sample size and saturation in PhD studies using qualitative interviews. Forum Qual Soc Res 2010 Sep;11(3):art 8.

Abbreviations

COREQ: Consolidated Criteria for Reporting Qualitative Research
EAM: electronic activity monitor

Edited by G Eysenbach; submitted 29.05.17; peer-reviewed by R Rudin; accepted 23.08.17; published 23.09.17.

Please cite as:
Bellicha A, Macé S, Oppert JM
Prescribing of Electronic Activity Monitors in Cardiometabolic Diseases: Qualitative Interview-Based Study
J Med Internet Res 2017;19(9):e328
URL: http://www.jmir.org/2017/9/e328/
doi:10.2196/jmir.8107
PMID:28947415

Original Paper

Social Media for the Dissemination of Cochrane Child Health Evidence: Evaluation Study

Michele P Dyson[1,2], PhD; Amanda S Newton[2], RN, PhD; Kassi Shave[1,2], MSc; Robin M Featherstone[1,2,3], MLIS; Denise Thomson[3], MSc, MBA; Aireen Wingert[1,2], MPH; Ricardo M Fernandes[3,4,5], MD, PhD; Lisa Hartling[1,2,3], MSc, PhD

[1]Alberta Research Centre for Health Evidence, University of Alberta, Edmonton, AB, Canada
[2]Department of Pediatrics, University of Alberta, Edmonton, AB, Canada
[3]Cochrane Child Health, University of Alberta, Edmonton, AB, Canada
[4]Faculty of Medicine, Universidade de Lisboa, Lisbon, Portugal
[5]Cochrane Portugal, Universidade de Lisboa, Lisbon, Portugal

Corresponding Author:
Michele P Dyson, PhD
Alberta Research Centre for Health Evidence
University of Alberta
4-474 Edmonton Clinic Health Academy, 11405-87 Avenue NW
Edmonton, AB, T6G 1C9
Canada
Phone: 1 780 492 1241
Fax: 1 780 248 5627
Email: mdyson@ualberta.ca

Abstract

Background: Health care providers value ready access to reliable synthesized information to support point-of-care decision making. Web-based communities, facilitated by the adoption of social media tools such as Facebook, Twitter, and YouTube, are increasingly being used for knowledge dissemination, bridging the gap between knowledge generation and synthesis and knowledge implementation.

Objective: Our objective was to implement and evaluate a structured social media strategy, using multiple platforms, to disseminate Cochrane Child Health evidence to health care providers caring for children.

Methods: Our social media strategy had three components: daily “tweets” using the Cochrane Child Health Twitter account, weekly WordPress blog posts, and a monthly journal club on Twitter (“tweet chat”). Each tweet, blog, and journal club shared Cochrane evidence on a child health topic. We evaluated the strategy through (1) Twitter and blog site analytics, (2) traceable link (Bitly) statistics, (3) Altmetric.com scores for promoted evidence, and (4) participant feedback. We also tracked the resources required to write the blog, tweet content, and manage the strategy.

Results: The 22-week social media strategy ran between November 2014 and April 2015. We created 25 blog posts, sent 585 tweets, and hosted 3 tweet chats. Monthly blog visits and views and Twitter account followers increased over time. During the study period, the blog received 2555 visitors and 3967 page views from a geographically diverse audience of health care providers, academics, and health care organizations. In total, 183 traceable Bitly links received 3463 clicks, and the Twitter account gained 469 new followers. The most visited and viewed blog posts included gastrointestinal topics (lactose avoidance), research on respiratory conditions (honey for cough and treatments for asthma), and maternal newborn care (skin-to-skin contact). On Twitter, popular topics were related to public health (vaccination) and pain management. We collected Altmetric.com scores for 61 studies promoted during the study period and recorded an average increase of 11 points. Research staff (n=3) contributed approximately 433 hours to promotion activities and planning (6.5 hours each per week) to implement the social media strategy, and study investigators reviewed all content (blog posts and tweets).

Conclusions: This study provides empirical evidence on the use of a coordinated social media strategy for the dissemination of evidence to professionals providing health services to children and youth. The results and lessons learned from our study provide guidance for future knowledge dissemination activities using social media tools.

(J Med Internet Res 2017;19(9):e308) doi:10.2196/jmir.7819

KEYWORDS

social media; translational medical research; health knowledge, attitudes, practice; pediatrics

Introduction

Advances in technology have markedly changed the way individuals can access and use information. The use of social media and Web 2.0 technologies is rapidly changing the health landscape, redefining the way health care providers connect professionally with colleagues and patients [1,2]. Social media tools such as Facebook, Twitter, and YouTube are increasingly being used by health care providers to access virtual communities where research evidence can be shared and exchanged [3-6]. Please see Multimedia Appendix 1 for a glossary of social media terms. The Web-based interface of social media transcends many traditional geographic barriers, creating the potential for connecting health care providers who might not otherwise interact [7]. Health care providers value ready access to highly synthesized and reliable information to support point-of-care decision making. The wide reach and accessibility of social media create opportunities for expanded dissemination of evidence among professional health care networks and ultimately increase opportunities for the uptake and implementation of evidence in practice [8].

Social media may have great potential for use as a rapid, accessible, and cost-effective strategy to disseminate knowledge to health care professionals [8]. Proponents of the use of social media for knowledge translation in health care point to three key features that make these tools highly effective: personalization, presentation, and participation [9]. The tailoring of content allows users to access and share information that is most valuable to them, whereas the versatility of social media creates numerous options for the presentation of information. The immediacy of social media also facilitates timely information sharing, and the availability of multiple formats (eg, blogging platforms, microblogging sites, and social networking sites) allows for flexible dissemination options, depending on the purpose of the tools and the preferences of the target population. The collaborative nature of social media allows for a meaningful contribution from all user groups [5]. Finally, social media can incorporate components of traditional knowledge translation interventions that have demonstrated effectiveness in changing health care providers' behavior, including the combination of didactic and interactive content in the distribution of educational materials [10] and endorsement by local opinion leaders [11].

The body of literature exploring social media and their utility in health care is rapidly growing; however, focus is primarily given to social media as a tool that patients can utilize to support their health and how social media can be used to enhance communication between patients and health care professionals [1]. A recent integrative review reported a modest level of evidence that a desire to gain and exchange knowledge is a primary motivator for social media use by health care professionals [12]. The evidence suggests that clinicians communicate via social media mostly within their discipline and that gaining access to new knowledge is an essential benefit of engagement with social networks and virtual communities [12]. Despite the potential for social media as a knowledge translation strategy in health care and the enthusiasm surrounding its use, there is a lack of empirical and longitudinal studies examining the effectiveness of using social media as a basis for a knowledge mobilization strategy aimed at health care professionals [13].

Cochrane (formerly known as the Cochrane Collaboration) is an international network of health researchers, professionals, and consumers who work together to synthesize high-quality, trusted evidence to enhance health care knowledge and decision making [14]. The members of Cochrane translate review evidence for different audiences using a variety of formats such as decision aids, plain language summaries of Cochrane systematic reviews, and podcasts [15]. Cochrane entities, including Cochrane Musculoskeletal, Cochrane Croatia, and Cochrane Schizophrenia, have experimented with the use of social media tools (Facebook and Twitter) to disseminate Cochrane summaries [15-17]. Between March 2013 and June 2014, Cochrane Croatia measured Facebook activity for a page sharing Croatian translations of Cochrane summaries and gained 1441 followers. Contributing to the empirical evidence of the effectiveness of social media dissemination, Cochrane Schizophrenia designed a randomized controlled trial to assess the impact of tweeting Cochrane summaries and showed a nearly threefold increase in Web visits to the tweeted Cochrane content [17].

Cochrane Child Health advocates for and facilitates the conduct of systematic reviews on child health topics [18]. To build on the exploratory use of social media for knowledge translation of Cochrane evidence, the objectives of this study were to implement and evaluate a structured social media strategy using multiple platforms to disseminate child health–relevant Cochrane systematic reviews and summaries to health care providers caring for children.

Methods

Summary of Promotional Activities

For a 22-week study period between November 3, 2014 and April 5, 2015, we promoted high-quality child health evidence to professionals providing health services to children and youth through Cochrane Child Health's social media presences. Our strategy comprised three key components: (1) a weekly blog post, (2) daily messages on Twitter ("tweets"), and (3) a monthly journal club hosted on Twitter as a "tweet chat." We collected data on Web traffic and user engagement through metrics provided by a series of Web-based analytics tools described below.

Identification of Evidence and Vetting of Content

Our selection of evidence to promote was guided by content contained in child health–relevant systematic reviews within the Cochrane Database of Systematic Reviews. Our efforts were informed by popular and topical Cochrane content, material included in *Evidence-Based Child Health: A Cochrane Review Journal* (eg, overviews of reviews and podcasts), and collaborations with other bloggers associated with Cochrane. Each week, we chose a new topic to profile as the focus of our blog posts and tweets. Our goal was to represent high-quality evidence across a spectrum of child-relevant issues to appeal to a diverse group of child health care providers. The blog posts were drafted by study investigators (MPD, ASN, DT, and LH), study staff (KS and AW), or content experts. All content (ie, blog posts and tweets) was reviewed by study investigators, and relevant content experts as needed, to ensure the accuracy of information shared.

Social Media Strategy

Our strategy focused on two commonly used social media tools: blogs and Twitter [19]. At the beginning of every week during the study period, we added a new post to our WordPress blog, summarizing the key messages from the Cochrane review or overview being featured that week (see Multimedia Appendix 2 for blog titles and topic categories). The blog posts were written with the goal of being succinct and written in plain language to facilitate uptake by our end users [20]. We also incorporated images and maximized the use of white space to increase visual appeal [21]. Along with our summary, we provided an appraisal of the evidence and links to the original research and supplementary material such as podcasts, Cochrane plain language summaries, and patient resources (see Multimedia Appendix 3 for a sample blog post).

For the remainder of each new week, we promoted our blog post and the evidence on Twitter (@Cochrane_Child). We published 21 tweets per week (3 tweets per day, 7 days per week). We used the Web-based scheduling tool Twuffer to ease daily resource demands and timed our tweets to be released in the morning, afternoon, and evening (local Mountain Standard Time). Common Twitter hashtags (keywords or phrases preceded by a hash symbol [#] to identify messages on specific topics) and handles (Twitter account names preceded by the @ symbol) for relevant pediatric interest groups were included in every tweet. The @Cochrane_Child Twitter account was monitored by a research staff member to respond to comments and engage with followers as applicable.

Over the course of the 22-week period, we hosted 3 journal clubs (tweet chats) on Twitter (see Multimedia Appendix 4 for journal club titles and topic categories). Each tweet chat was an hour long and took place at a prescheduled date and time. We promoted the meetings with blog posts and tweets. For our tweet chat meetings, we also recruited clinical experts known to the research team on the chosen topics to participate. Our expert collaborators helped us identify key discussion points related to the quality and applicability of the evidence. We used these discussion points to help facilitate the tweet chat and also to promote the event. The tweet chats were intended to be informal, allowing for discussion among and questions from participants. However, we prepared tweets and a rough script for the meeting to maintain a consistent format and to ensure that all key points would be addressed. Journal club meetings were recorded and archived using the Web-based tool Storify. We promoted the archived tweet chats on the WordPress blog and Twitter.

Evaluation

Blog

Web traffic to and user engagement with the blog were measured using the built-in analytics program in WordPress. We tracked the numbers of page views and visitors per day, sources of site referrals, and the geographic spread of our visitors. Blog posts were open for comments, which we collected and analyzed.

Twitter

Twitter analytics and Twitonomy were used to obtain detailed metrics for our tweets, including engagement (number of times users interacted with our tweets), impressions (potential number of times users viewed our tweets), and the number of retweets (number of times users shared our tweets with their followers), clicks, favorites, and followers. We tracked specific tweets that received the most attention and classified these using the Cochrane review group structure to clarify the relevant clinical area. We also collected data on the number of followers the @Cochrane_Child Twitter account gained during the promotion and their behavior over the study period. Where possible, we extracted descriptive data (eg, profession, affiliation, and self-reported interests) on our followers from their public profiles.

Journal Clubs

Following each journal club, we asked participants to complete a brief survey using Google Forms. The survey was distributed via Twitter and comprised 8 multiple-choice and free-text questions (available in Multimedia Appendix 5). Ethics approval for the survey was obtained from the Health Research Ethics Board at the University of Alberta. For the tweet chat journal clubs, we also recorded the number of views each of the archived meetings received on Storify.

Accessing the Original Publications

To evaluate the impact of our social media strategy on the frequency that the original publications were accessed, we used traceable links generated by Bitly and alternative social media metrics (altmetrics) through Altmetric.com scores. For all hyperlinks that we posted on our blog and Twitter, we used shortened, unique Web addresses (URLs) created by Bitly. The Bitly account collected data on the number of clicks each link received, allowing us to directly measure audience interaction with our social media posts. Altmetrics measure Web-based attention based on how far scholarly content travels through the social Web and encompass reflections of both the quantity and the quality of attention received [22]. The Altmetric.com score is an automatically calculated weighted count of all attention a publication has received, based on the volume, sources, and authors [23]. The score increases as more people mention the publication; however, the amount by which the score will increase varies based on the source of the mention (eg, the score will increase more if the source of the mention was a newspaper vs a tweet), self-promotion by authors, potential bias toward a journal or a publisher, and whether the content is being shared directly with its intended audience [23]. A decrease in the score may occur in unusual cases because of a fluctuation. A fluctuation can result from a tweet being removed by the original tweeter, a Twitter account being deemed “biased” according to an Altmetric.com moderator, or because of a change made to the algorithm being used to calculate the score [24]. For each publication we shared throughout the study, we collected Altmetric.com data before, during, and after the promotion effort. We attempted to collect data on article downloads for the promoted Cochrane systematic reviews from the publisher, but only the annual numbers of downloads and page views were available. We were unable to obtain information specific to our study period.

Resource Implications

Using an internal time log, we tracked data on the staff time dedicated to this project, including the upkeep of the social media accounts, identifying content, writing and publishing blog posts and tweets, and participating in journal club meetings. The data available do not reflect the time committed by the study investigators and other content experts.

Analysis

We used descriptive statistics, including numbers, frequencies, and means, to analyze quantitative data and content analysis to analyze qualitative data where applicable [25]. We collected data on our blog site, Twitter account, and Altmetric.com scores at the outset to establish a baseline of usage to compare data collected during the promotion; we continued to collect data for click counts and page views for 1 month beyond the promotion to allow our messages and posts time to be viewed and shared on social media sites. Popularity of blogs was determined by the number of visitors and views, and popularity of tweets was determined by the number of retweets, favorites, engagements, impressions, and URL click counts. The results were graphically displayed.

Results

Summary of Web-Based Attention Received

Over the study period, we shared 25 blog posts (22 topic posts and 3 journal club announcements; Multimedia Appendices 2 and 3), published 585 tweets (434 promoting the blog posts, 36 promoting the journal clubs, and 115 during journal club meetings), and hosted 3 journal club sessions. Overall, the blog received a total of 2555 visitors and 3967 page views, and the Twitter account gained 469 new followers from a baseline of 596 followers to 1065 followers at study completion (representing a 79% increase, 469/5.96). The degree of monthly engagement with both the blog and Twitter account increased over time (Figure 1).

Web Traffic and User Engagement

The following data were collected during the period October 6, 2014 to April 24, 2015, and data collection overlaps the study period by approximately 1 month on either end. Most views of the blog (50%, 1996/39.67) originated via referrals from Twitter, followed by Facebook (16%, 640/39.67). Facebook referrals were likely mainly from the accounts managed by Cochrane and Cochrane Canada. Also, one of our research collaborators (RF) had linked Facebook and Twitter accounts, which mirrored posts composed on each site. Of the total views, 1.4% (56/39.67) (n=56) originated from the Cochrane Child Health website. Geographically, views were mainly from Canada (29%, 1181/39.67) and the United Kingdom (28%, 1120/39.67); however, the remaining views were distributed among an international audience (North America, excluding Canada: 24%, 946/39.67; Europe: 23%, 897/39.67; Australia: 5%, 207/39.67; Africa: 3%, 117/39.67; South America: 4%, 175/39.67; and Asia: 1%, 44/39.67).

The most popular blog topics were (1) “Honey: An effective cough remedy for kids?” (n=241 visitors; n=359 page views); (2) “To wheeze or not to wheeze” (long-acting beta2-agonists [LABAs] for asthma) (n=216 visitors; n=340 page views); and (3) “Lactose avoidance: Worthwhile for reducing duration of diarrhea in kids?” (n=215 visitors; n=304 page views). These 3 blog posts were published in consecutive weeks (study weeks 16-18), and “LABAs for asthma” was the topic of our second journal club (during study week 17). Overall, 3 posts generated comments from viewers: “The power of touch: Skin-to-skin contact and kangaroo mother care for newborns” (link to evidence regarding skin-to-skin contact and pain management), “To wheeze or not to wheeze” (clinical comment regarding use of LABAs and inhaled corticosteroids for asthma), and “Children and youth with obesity—A growing global epidemic” (notification that our post had been reblogged).

Our Twitter account attracted mainly individuals who identified as health care providers (n=206) and academics (n=107). We also gained some attention from individuals representing organizations such as hospitals, nonprofit organizations, and professional associations (n=65). Self-reported interests often included child health and care (n=120), with additional categories including health evidence and resources (eg, evidence-based practice and knowledge translation; n=29) and social justice, advocacy, and policy (n=27). The classifications reported were not mutually exclusive.

Across clinical topics, our tweets were most frequently in the areas of acute respiratory infections (ARI; n=139), neonatology (n=43), and airways (n=42). Twitter followers were highly engaged with tweets related to public health (mean 4.5 retweets per tweet and 1.7 favorites per tweet) and pain (mean 3.6 retweets per tweet and 1.6 favorites per tweet). These results are summarized in Table 1. For individual tweets, the highest numbers of retweets were in the areas of ARI (antivirals for influenza; n=17), public health (obesity prevention programs; n=13), and airways and ARI (promotion for asthma journal club and measles-mumps-rubella [MMR] vaccination; n=12 each). Impressions were highest for 2 tweets about the MMR vaccine (n=3151 and 2983) and public health (n=2374), engagement was greatest for a tweet related to sucrose/glucose for infant pain (n=51), and the largest number of URL clicks were in tweets related to ARI (antibiotics for sore throat, n=28; bronchiolitis, n=20) and neonatology (procedural pain, n=25). See Table 2 for details.

Figure 1. Engagement with the Cochrane Child Health blog and @Cochrane_Child Twitter account over time.

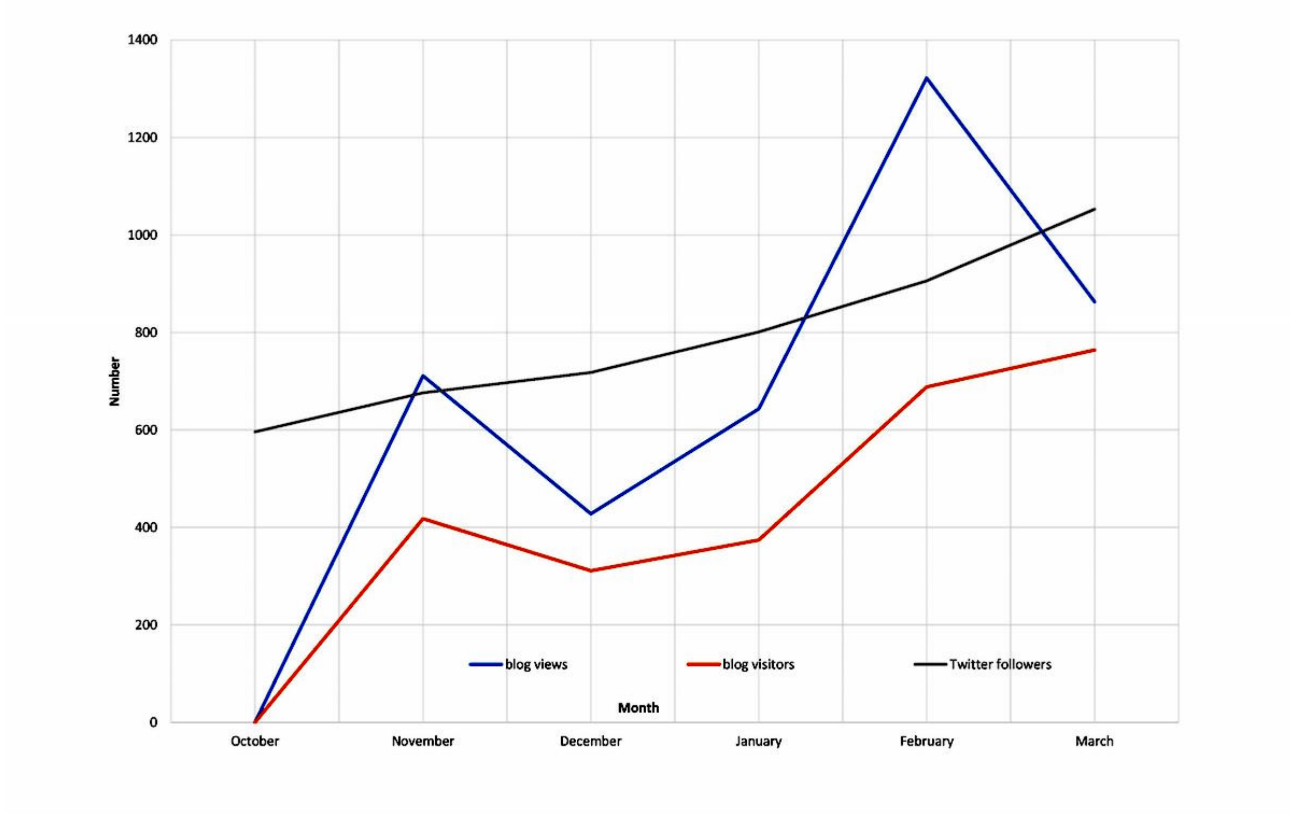


Table 1. Summary of Twitter activity by tweet topic.

Topic	Total tweets (N=4.34), n (%)	Total retweets (N=10.28), n (%)	Mean retweets per tweet	Total favorites (N=3.57), n (%)	Mean favorites per tweet
Acute respiratory infections	139 (32)	318 (31)	2.3	126 (35)	0.91
Neonatology	43 (10)	92 (9)	2.1	23 (6)	0.53
Airways	42 (10)	80 (8)	1.9	47 (13)	1.1
Anxiety, depression, and neurosis	22 (5)	43 (4)	2	11 (3)	0.5
Cystic fibrosis and genetic disorders	22 (5)	19 (2)	0.86	4 (1)	0.18
Fertility regulation	21 (4.8)	36 (4)	1.7	9 (3)	0.43
Infectious diseases	21 (4.8)	45 (4)	2.1	16 (5)	0.76
Injuries	21 (4.8)	35 (3)	1.7	4 (1)	0.19
Pain, palliative and supportive care	21 (4.8)	76 (7)	3.6	34 (10)	1.6
Psychosocial and learning problems	21 (4.8)	43 (4)	2	8 (2)	0.38
Skin disorders	21 (4.8)	18 (2)	0.86	12 (3)	0.57
Cancer	20 (4.6)	134 (13)	3.4	29 (8)	0.73
Public health	20 (4.6)	89 (9)	4.5	34 (10)	1.7
Total	434	1028	2.23	357	0.74

Table 2. Summary and ranking of four most popular tweets by specific metrics.

Tweet topic	Tweet content	Retweets, n (rank)	Favorites, n (rank)	Impressions[a], n (rank)	Engagement[b], n (rank)	URL clicks, n (rank)
Acute respiratory infections	Do antivirals work for #flu? @cochranecollab evidence concludes they're not as effective as we previously thought	17 (#1)	3 (-)[c]	2175 (#4)	50 (#2)	7 (-)
Public health	#systematicreview of 37 studies w/27,964 children shows #obesity prevention programs reduce adiposity	13 (#2)	6 (#3)	1400 (-)	44 (-)	12 (-)
Journal club promotion	Passionate about #childhealth evidence? Join the #CochraneChild tweet chat Feb 25 @ 2 pm MST (9 pm GMT)	12 (#3)	6 (#2)	1974 (-)	45 (-)	8 (-)
Acute respiratory infections	Measles is making a comeback | #itsasmallworldafterall, so get vaccinated!	12 (#3)	1 (-)	3151 (#1)	32 (-)	3 (-)
Pain, palliative and supportive care	Pain relief important part of caring for kids | we need to address knowing-doing gap. #painevidence	11 (#4)	4 (-)	1409 (-)	45 (-)	13 (-)
Neonatology	No more research needed on sucrose/glucose for infants pain | Now we need to put knowledge into action!	11 (#4)	5 (#4)	1690	51 (#1)	25 (#2)
Acute respiratory infections	Blogging #childhealth evidence | This week: Cochrane meets controversy: Vaccines for measles, mumps, & rubella	11 (#4)	4 (-)	2983 (#2)	43 (-)	13 (-)
Public health	Blogging #childhealth evidence | This week: Policies & strategies for preventing childhood #obesity	11 (#4)	5 (#4)	2347 (#3)	39 (-)	6 (-)
Acute respiratory infections	Coughing kids? #honey better than placebo or diphenhydramine for improving sleep in children (and parents)	10 (-)	6 (#1)	1149 (-)	46 (-)	12 (-)
Acute respiratory infections	1 out of 3 babies will get #bronchiolitis in their 1st year | Despite its prevalence, clinical practice varies	7 (-)	3 (-)	823 (-)	49 (#3)	20 (#3)
Infectious diseases	Oral rehydration? Lactose avoidance? Both? | @giordanopg talks about how to treat acute diarrhea in kids	7 (-)	5 (-)	1186 (-)	49 (#3)	15 (-)
Acute respiratory infections	Are antibiotics over-prescribed for sore throat? @UKCochraneCentr | #childhealth evidence	6 (-)	5 (-)	1525 (-)	47 (#4)	28 (#1)

[a] Impressions reflect the number of times a user is served a tweet in timelines or search results.

[b] Engagement reflects the total number of times a user interacted with a tweet.

[c] (-) indicates that the tweet did not rank in the top four.

Tweet chats had limited numbers of active Web-based participants but gained usage through a Web-based repository where the journal clubs could be accessed after the meeting. The archived tweet chats on Storify received between 37 and 57 views each: January—bronchiolitis (n=57); February—asthma (n=22); and March—obesity prevention (n=37). Three participants responded to the surveys administered after each journal club: one in response to the bronchiolitis journal club and two in response to the asthma journal club. All 3 participants were physicians, with 1 physician being interested in Twitter generally and the other 2 physicians being specifically interested in the content (asthma). The perceived benefits of hosting a journal club on Twitter included it being a useful tool to check understanding of the subject matter and gain new ideas, as well as providing a level playing field for everyone interested in participating. A suggestion to improve the format of the journal club was to include key images with the tweets to add to the discussion.

Accessing the Original Publications

Over the course of the study, we created 183 customized, traceable Bitly links, which received 3463 clicks. Just over half of these clicks (55%, 1892/34.63) were directed to the Cochrane Child Health blog home page. As our new posts always appeared on the blog's home page, we consistently promoted the home page link on social media channels. A considerable proportion of clicks (14%, 468/34.63) were related to the journal clubs, including the studies being discussed (7%, 232/34.63), links to the announcements (4%, 141/34.63), and archived discussions (2%, 61/34.63). The Cochrane review on interventions for preventing obesity in children was the most highly accessed (n=93 clicks; see Table 3; [26]). Other commonly accessed studies included those on oral antihistamine-decongestant-analgesic combinations for the common cold (n=99) and procedural pain in children (n=59; [27,28]).

We collected Altmetric.com data for 61 studies promoted during the study period (Multimedia Appendix 6). The mean change in score was an increase of 11 points (median: 5; range: −1 to

73). Usage for our study corresponded to an average of 10 clicks on the Bitly links to the studies that we promoted (median: 3; range: 0-97) and a mean change in Altmetric.com score per Bitly click of 3 (median: 1.5; range: −0.25 to 37). Most attention for these studies came from Twitter (n=2229 tweeters), Mendeley (n=981 readers), and Facebook (n=400 timelines). The topics of studies that experienced the greatest Altmetric.com score increase included the following: (1) neuraminidase inhibitors for preventing and treating influenza, (2) vitamin C for the common cold, and (3) zinc for the common cold (see Table 4). Of our journal club studies, the Altmetric.com scores increased by 36 points (range: 104-139) for the review on interventions for preventing obesity for children [26], 17 points (range: 0-17) for the review on LABAs for asthma [29], and 11 points (range: 0-11) for the overview on the treatment of bronchiolitis [30]. These scores placed the reviews and overviews in the 99th, 95th, and 88th percentiles, respectively, for Altmetric.com scores of studies of the same age and published in any journal.

Table 3. Most commonly accessed links.

Rank	Resource title link	Link type	Clicks
#1	Cochrane Child Health blog home page	Blog home page	1892
#2	Oral antihistamine-decongestant-analgesic combinations for the common cold	Cochrane systematic reviews	99
#3	The Cochrane Library and procedural pain in children: An overview of reviews	Cochrane overview of systematic reviews	97
#4	Interventions for preventing obesity in children	Cochrane systematic reviews	93
#5	The Cochrane Library and the treatment of bronchiolitis in children: An overview of reviews	Cochrane overview of systematic reviews	77
#6	Journal club announcement: LABAs for asthma	Journal club announcement	65
#7	The Cochrane Library and safety of regular long-acting beta2-agonists in children with asthma: An overview of reviews	Cochrane overview of systematic reviews	62
#8	The Cochrane Library and procedural pain in children: An overview of reviews [podcast]	Podcast	59
#9	Journal club announcement: Evidence for treatment of bronchiolitis	Journal club announcement	38
#9	Journal club announcement: Obesity prevention	Journal club announcement	38
#10	The Cochrane Library and the treatment of sore throat in children and adolescents: An overview of reviews	Cochrane overview of systematic reviews	37
#10	Does this patient have strep throat? The rational clinical examination	External evidence-based medicine resource	37
#10	Honey for acute cough in children	Cochrane systematic reviews	37

Table 4. Top Altmetric.com score growth among promoted studies.

Article title	Journal	Total score increase, points (% increase)	Baseline score	Final score
Neuraminidase inhibitors for preventing and treating influenza in healthy adults and children	Cochrane Database of Systematic Reviews (CDSR)	73 (23)	312	385
Vitamin C for preventing and treating the common cold	CDSR	53 (18)	295	348
Zinc for the common cold	CDSR	40 (15)	271	311
Interventions for preventing obesity in children	CDSR	36 (35)	104	140
Honey for acute cough in children	CDSR	35 (97)	36	71

Resource Implications

Not including the time invested by the study investigators to write and review blog posts and tweets and to plan and participate in the journal club meetings, research staff (n=3) contributed approximately 433 hours to the project (approximately 6.5 hours each per study week). We involved a research librarian, a project coordinator, and a graduate student throughout the project; 2 undergraduate summer students were also briefly involved. These members of the team led the coordination of our Twitter activity and logistics of managing the blog and also contributed to drafting blog posts.

Discussion

Empirical Evidence of Social Media for Knowledge Dissemination

This study provides empirical evidence on the potential impact of social media activities for knowledge translation in the health sciences. We implemented a structured 22-week social media strategy involving a coordinated approach using commonly accessed social media platforms (ie, WordPress blog and Twitter) and demonstrated that engagement with our blog and Twitter account increased steadily over time and was geographically diverse. However, we found our approach was resource intensive and required the involvement of several content experts.

Social media platforms have been widely explored in the context of facilitating communication and improving knowledge among health care professionals [4]. Our findings demonstrate that a coordinated social media strategy may be an effective approach for sharing health evidence among a geographically diverse audience of health care providers, academics, and health care organizations. Although our promotional activities originated in Canada, we also attracted attention across North America, Europe, Australia, Africa, Asia, and South America. Similarly, in a study conducted by Cochrane Croatia utilizing Facebook as a dissemination tool for Cochrane summaries, the intended audience was initially nationally focused but ultimately grew to include followers worldwide [16]. Ultimately, the reach of social media can far surpass intended immediate audiences, which reinforces the potential utility of social media tools for extending the global reach of health research.

Although social media can be an effective means for broadly sharing health research, the process is resource intensive and requires careful planning. To realize our social media strategy, the research staff (n=3) contributed approximately 433 hours to promotion activities and planning (6.5 hours each per week). In addition to research staff hours, each week, the study investigators were involved in reviewing all content (blog posts and tweets). Content experts provided a quote on the relevant review and, for journal clubs, moderated discussion for 1 hour. Investigators and content experts were involved to ensure information accuracy; however, this process was resource intensive and involved a trade-off in terms of effort for yield. Moreover, engaging content experts can be challenging because of the time required without formal acknowledgment of their contributions (eg, through academic recognition).

Strengths and Limitations

Previous research on disseminating Cochrane evidence through social media has primarily utilized single platforms such as Facebook [16] or Twitter [17] and has minimally utilized Web analytics (eg, page views or number of followers) to provide a quantitative impact assessment. To the best of our knowledge, this is the first study where a Cochrane group or knowledge translation program has evaluated a social media strategy using coordinated platforms (eg, Twitter and WordPress blogging) and a range of analytics (altmetrics, click counts, page views, site visits, and engagements) from different sources (Altmetric.com, Twitter analytics, Bitly, and WordPress analytics) to create a more complete and quantifiable usage picture. Recently, Cochrane review groups have highlighted their support for journal club activities [15], yet little research describing and evaluating the process of conducting journal clubs via Twitter (tweet chats) has been conducted. Our social media strategy provides insight with regard to this gap in the published literature through the development, implementation, and quantitative assessment of 3 tweet chats. Additionally, the quantification of required resources and staff time is an additional contribution of this study to the literature on social media for knowledge dissemination and will allow other research teams to estimate the time from staff and others needed to undertake such a project. This information is useful to a range of users, including those designing social media strategies, those interested in disseminating health information in general, and those interested in promoting specific items (eg, authors or publishers wishing to promote specific items).

Limitations of this study include a variable range of reporting periods for the statistics. Some statistics from free Web-based tools (ie, Bitly) were only available for a limited period of 30 days. In the case of Bitly click counts on our traceable URLs, we recorded data from the last 30 days of the promotion. In other cases, available statistics covered a longer period (ie, WordPress) than the 22-week promotion. WordPress blog site statistics were available for “all time” and extended back to when the blog was first created on October 6, 2014. The extended period of data collection may have contributed to an overestimate to usage than what we can directly attribute to our promotion. However, as the blog had very little activity before the promotion started, we do not consider that this extended reporting period contributed to a significant increase in usage. In the case of the journal clubs, we were unable to measure the number of passive Twitter users who viewed the tweet chat during the events. Recorded statistics for the Web archive of the journal clubs extended long past the time of the meetings (recorded in December 2016). Although inconsistent with the rest of our data collection activities, the tweet chat archive suggests continuing use of the materials, and we felt this information was worth collecting. A further limitation of the freely available analytics tools was an inability to exclude statistics from internal use. Our project team was aware of this problem and made every effort to limit the amount of site testing and internal clicks on our traceable links. There is, however, a possibility that some of our recorded numbers represent usage from our own team. Our study was also limited by the method in which Altmetric.com scores are informed. Not only do the Altmetric.com data capture our influence on evidence uptake, but they also encompass the impact of other sources during the same time frame. Due to this, we cannot distinguish between the impact of our social media promotion on article Altmetric.com scores and that of other sources. Finally, a greater potential challenge of this study, and to all dissemination research related to health evidence, is our inability to attribute knowledge sharing to behavior change and improvement in health outcomes. Our use of proxy measures allows for an overview of potential use of evidence; however, future research is needed to determine the health impact of knowledge-sharing activities, as well as methods to gather this information.

Lessons Learned and Future Directions

We have identified several lessons learned (see Textbox 1) from our study that we believe are applicable to individuals interested in utilizing social media as a knowledge dissemination strategy across a variety of academic and professional fields. Foremost, the awareness of the amount of time and effort required to manage the social media strategy, including identifying and creating content, is paramount. In this study, our recorded time was limited to what was invested by research staff; investigators and content experts (eg, clinicians and guest bloggers) invested additional time weekly that was not captured in our results. Also important to consider are the specific objectives or purposes of the social media strategy. We focused our efforts almost entirely on dissemination and thus generated very minimal interactivity, which in many cases is one of the intended advantages of social media. Further research evaluating processes and outputs related to purposeful approaches to increased interactivity is required. However, increasing interactivity will require greater investment in terms of resources and time. This, in part, may be mediated by pursuing focused topic areas and utilizing existing established networks. Finally, considering the scope of topics included in the social media strategy and the associated resource implications of the chosen scope would be of benefit. In this study, we found our broad focus (ie, child health evidence) and our approach of changing specific topics each week to be challenging in terms of finding and engaging numerous content experts. Social media strategies within specific content areas may be more effective at developing a social network or engaging within an existing Web-based community, as well as activating relevant opinion leaders. Subsequently, a more content-specific approach may have a greater impact on information uptake in practice.

Priorities for future research include identifying metrics to assess the impact of social media and a given social media strategy's intended effects. Furthermore, conducting research focusing on a better understanding of the experiences and motivations of social media users and the uptake of disseminated information in practice would be of benefit.

Textbox 1. Lessons learned.

- Consider allocation of time and budgeting resources for all personnel before commencing the social media strategy
- Consider the specific objectives of the social media strategy
- Consider the scope of topics to be covered in the social media strategy (eg, narrow or broad)
- Utilize existing established networks when available and feasible

Conclusions

Our coordinated knowledge dissemination strategy allowed us to gather empirical evidence on how social media can be used to share research evidence with an audience of child health professionals. We increased Web-based followers for our social media presences and, using proxy measures, observed an increase in access to the evidence we were promoting. Time tracked by our team members provides an estimate for researchers planning to undertake a similar promotion using social media tools. The range of analytics included in this study contributes to our understanding of how to assess the reach and impact of knowledge dissemination activities via social media. Our methods of using coordinated activities via multiple social media platforms expand on existing dissemination practice and explore numerous opportunities to enhance health research promotion.

Acknowledgments

The authors would like to acknowledge Dr Ian Sinha and Dr Geoff Ball for generously donating their time to moderate journal clubs for this study. We would also like to thank Dr Brian Rowe, Dr Antonia Stang, Dr Joan Robinson, and Dr Samina Ali for providing expertise and assistance in producing content for the blog and Dr David King, Dr Bob Phillips, and Sarah Chapman for giving advice on moderating journal clubs and suggesting potential metrics for measurement. Finally, we would like to thank Sarah Chapman, Dr Giordano Pérez-Gaxiola, and Tess Griffin and her family for contributing guest posts during this study. This work was supported by Alberta Innovates—Health Solutions, grant number 201300653. LH and ASN are supported, in part, by Canadian Institutes of Health Research New Investigator Awards. The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Glossary of social media terms.

[PDF File (Adobe PDF File), 22KB - jmir_v19i9e308_app1.pdf]

Multimedia Appendix 2

WordPress blog post topics by week.

[PDF File (Adobe PDF File), 30KB - jmir_v19i9e308_app2.pdf]

Multimedia Appendix 3

A sample post from the Cochrane Child Health WordPress blog.

[PDF File (Adobe PDF File), 101KB - jmir_v19i9e308_app3.pdf]

Multimedia Appendix 4

Journal club topics by week.

[PDF File (Adobe PDF File), 22KB - jmir_v19i9e308_app4.pdf]

Multimedia Appendix 5

Journal club survey questions.

[PDF File (Adobe PDF File), 20KB - jmir_v19i9e308_app5.pdf]

Multimedia Appendix 6

Altmetric.com scores for all promoted articles.

[PDF File (Adobe PDF File), 53KB - jmir_v19i9e308_app6.pdf]

References

1. Frisch NC, Borycki EM, Mickelson G, Atherton P, Novak-Lauscher H, Hooker D, et al. Use of social media and web 2.0 technologies to increase knowledge and skills of british columbia nurses. NI 2012 2012;2012:117 [FREE Full text] [Medline: 24199067]
2. Hamm MP, Shulhan J, Williams G, Milne A, Scott SD, Hartling L. A systematic review of the use and effectiveness of social media in child health. BMC Pediatr 2014;14:138 [FREE Full text] [doi: 10.1186/1471-2431-14-138] [Medline: 24886048]
3. Giustini D. How Web 2.0 is changing medicine. BMJ 2006 Dec 23;333(7582):1283-1284 [FREE Full text] [doi: 10.1136/bmj.39062.555405.80] [Medline: 17185707]
4. Hamm MP, Chisholm A, Shulhan J, Milne A, Scott SD, Klassen TP, et al. Social media use by health care professionals and trainees: a scoping review. Acad Med 2013 Sep;88(9):1376-1383. [doi: 10.1097/ACM.0b013e31829eb91c] [Medline: 23887004]
5. 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;15(4):e85 [FREE Full text] [doi: 10.2196/jmir.1933] [Medline: 23615206]
6. Nordqvist C, Hanberger L, Timpka T, Nordfeldt S. Health professionals' attitudes towards using a web 2.0 portal for child and adolescent diabetes care: qualitative study. J Med Internet Res 2009;11(2):e12 [FREE Full text] [doi: 10.2196/jmir.1152] [Medline: 19403464]
7. Harris JK, Choucair B, Maier RC, Jolani N, Bernhardt JM. Are public health organizations tweeting to the choir? understanding local health department twitter followership. J Med Internet Res 2014 Feb 26;16(2):e31 [FREE Full text] [doi: 10.2196/jmir.2972] [Medline: 24571914]
8. Tunnecliff J, Ilic D, Morgan P, Keating J, Gaida JE, Clearihan L, et al. The acceptability among health researchers and clinicians of social media to translate research evidence to clinical practice: mixed-methods survey and interview study. J Med Internet Res 2015 May 20;17(5):e119 [FREE Full text] [doi: 10.2196/jmir.4347] [Medline: 25995192]
9. CDC. 2011. The health communicator's social media toolkit URL: http://www.cdc.gov/healthcommunication/ToolsTemplates/SocialMediaToolkit_BM.pdf [WebCite Cache ID 6pS7A1uSM]
10. Forsetlund L, Bjørndal A, Rashidian A, Jamtvedt G, O'Brien MA, Wolf F, et al. Continuing education meetings and workshops: effects on professional practice and health care outcomes. Cochrane Database Syst Rev 2009(2):CD003030. [doi: 10.1002/14651858.CD003030.pub2] [Medline: 19370580]
11. Flodgren G, Parmelli E, Doumit G, Gattellari M, O'Brien MA, Grimshaw J, et al. Local opinion leaders: effects on professional practice and health care outcomes. Cochrane Database Syst Rev 2011 Aug 10(8):CD000125 [FREE Full text] [doi: 10.1002/14651858.CD000125.pub4] [Medline: 21833939]
12. Rolls K, Hansen M, Jackson D, Elliott D. How health care professionals use social media to create virtual communities: an integrative review. J Med Internet Res 2016 Jun 16;18(6):e166 [FREE Full text] [doi: 10.2196/jmir.5312] [Medline: 27328967]
13. Schnitzler K, Davies N, Ross F, Harris R. Using Twitter™ to drive research impact: a discussion of strategies, opportunities and challenges. Int J Nurs Stud 2016 Jul;59:15-26. [doi: 10.1016/j.ijnurstu.2016.02.004] [Medline: 27222446]
14. Cochrane. About us URL: http://www.cochrane.org/about-us [accessed 2017-04-03] [WebCite Cache ID 6pS7SpFst]

15. Rader T, Pardo PJ, Stacey D, Ghogomu E, Maxwell LJ, Welch VA, et al. Update of strategies to translate evidence from cochrane musculoskeletal group systematic reviews for use by various audiences. J Rheumatol 2014 Feb;41(2):206-215. [doi: 10.3899/jrheum.121307] [Medline: 24293571]
16. Puljak L. Using social media for knowledge translation, promotion of evidence-based medicine and high-quality information on health. J Evid Based Med 2016 Dec 02;9(1):4-7. [doi: 10.1111/jebm.12175] [Medline: 26372327]
17. Adams CE, Jayaram M, Bodart AYM, Sampson S, Zhao S, Montgomery AA. Tweeting links to cochrane schizophrenia group reviews: a randomised controlled trial. BMJ Open 2016 Mar 08;6(3):e010509 [FREE Full text] [doi: 10.1136/bmjopen-2015-010509] [Medline: 26956164]
18. Cochrane. About us URL: http://childhealth.cochrane.org/about-us [accessed 2017-04-03] [WebCite Cache ID 6pS7ZZNSr]
19. Bik HM, Goldstein MC. An introduction to social media for scientists. PLoS Biol 2013;11(4):e1001535 [FREE Full text] [doi: 10.1371/journal.pbio.1001535] [Medline: 23630451]
20. Stableford S, Mettger W. Plain language: a strategic response to the health literacy challenge. J Public Health Policy 2007;28(1):71-93. [doi: 10.1057/palgrave.jphp.3200102] [Medline: 17363939]
21. Houts PS, Doak CC, Doak LG, Loscalzo MJ. The role of pictures in improving health communication: a review of research on attention, comprehension, recall, and adherence. Patient Educ Couns 2006 May;61(2):173-190. [doi: 10.1016/j.pec.2005.05.004] [Medline: 16122896]
22. Galligan F, Dyas-Correia S. Altmetrics: rethinking the way we measure. Serials Review 2013 Mar;39(1):56-61. [doi: 10.1016/j.serrev.2013.01.003]
23. Altmetric. 2017. The donut and altmetric attention score URL: https://www.altmetric.com/about-our-data/the-donut-and-score/ [accessed 2017-04-03] [WebCite Cache ID 6pS896dO8]
24. Altmetric. 2016. Why has the altmetric attention score for my paper gone down? URL: https://help.altmetric.com/support/solutions/articles/6000060969-how-is-the-altmetric-attention-score-calculated- [WebCite Cache ID 6pS8KchZD]
25. Hsieh H, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res 2005 Nov;15(9):1277-1288. [doi: 10.1177/1049732305276687] [Medline: 16204405]
26. Waters E, de SA, Hall BJ, Brown T, Campbell KJ, Gao Y, et al. Interventions for preventing obesity in children. Cochrane Database Syst Rev 2011(12):CD001871. [doi: 10.1002/14651858.CD001871.pub3] [Medline: 22161367]
27. De SA, van DM, Kumar AA, Lesslar O, Skrt A. Oral antihistamine-decongestant-analgesic combinations for the common cold. Cochrane Database Syst Rev 2012 Feb 15(2):CD004976. [doi: 10.1002/14651858.CD004976.pub3] [Medline: 22336807]
28. Curtis S, Wingert A, Ali S. and procedural pain in children: an overview of reviews. Evid.-Based Child Health 2012 Sep 06;7(5):1363-1399. [doi: 10.1002/ebch.1864]
29. Zhang L, Prietsch SO, Ducharme FM. Inhaled corticosteroids in children with persistent asthma: effects on growth. Cochrane Database Syst Rev 2014 Jul 17(7):CD009471. [doi: 10.1002/14651858.CD009471.pub2] [Medline: 25030198]
30. Bialy L, Foisy M, Smith M, Fernandes RM. The cochrane library and the treatment of bronchiolitis in children: an overview of reviews. Evid-Based Child Health 2011 Jan 26;6(1):258-275. [doi: 10.1002/ebch.673]

Abbreviations

altmetrics: alternative social media metrics
ARI: acute respiratory infection
LABA: long-acting beta2-agonist
MMR: Measles-Mumps-Rubella

Edited by G Eysenbach; submitted 03.04.17; peer-reviewed by K Rolls, P Caldwell; comments to author 18.05.17; revised version received 23.06.17; accepted 10.07.17; published 01.09.17.

Please cite as:
Dyson MP, Newton AS, Shave K, Featherstone RM, Thomson D, Wingert A, Fernandes RM, Hartling L
Social Media for the Dissemination of Cochrane Child Health Evidence: Evaluation Study
J Med Internet Res 2017;19(9):e308
URL: http://www.jmir.org/2017/9/e308/
doi:10.2196/jmir.7819
PMID:28864427

Original Paper

Cardiac Patients' Experiences and Perceptions of Social Media: Mixed-Methods Study

Stephanie R Partridge[1], PhD, APD; Anne C Grunseit[1], PhD; Patrick Gallagher[1], LLB; Becky Freeman[1], PhD; Blythe J O'Hara[1], PhD; Lis Neubeck[2,3,4], PhD; Sarah Due[5], RN, CNS, MPH; Glenn Paull[5], RN, MN, CNC; Ding Ding[1], PhD; Adrian Bauman[1], PhD, FAFPHM; Philayrath Phongsavan[1], PhD; Kellie Roach[6], RN, CNC; Leonie Sadler[7], RN, CNC; Helen Glinatsis[8], RN, CNC; Robyn Gallagher[3], RN, MN, BA, PhD

[1]Prevention Research Collaboration, Sydney School of Public Health, Charles Perkins Centre, The University of Sydney, Sydney, Australia
[2]School of Health and Social Care, Edinburgh Napier University, Edinburgh, United Kingdom
[3]Sydney Nursing School, Charles Perkins Centre, University of Sydney, Sydney, Australia
[4]School of Nursing and Midwifery, Faculty of Medicine, Nursing, and Health Sciences, Flinders University, Adelaide, Australia
[5]St. George Hospital, South Eastern Sydney Local Health District, Sydney, Australia
[6]Ryde Hospital, Northern Sydney Local Health District, Sydney, Australia
[7]Manly Hospital, Northern Sydney Local Health District, Sydney, Australia
[8]Royal North Shore Hospital, Northern Sydney Local Health District, Sydney, Australia

Corresponding Author:
Robyn Gallagher, RN, MN, BA, PhD
Sydney Nursing School
Charles Perkins Centre
University of Sydney
Room 2210, Level 2, Building D17
Sydney,
Australia
Phone: 61 286270279
Email: robyn.gallagher@sydney.edu.au

Abstract

Background: Traditional in-person cardiac rehabilitation has substantial benefits for cardiac patients, which are offset by poor attendance. The rapid increase in social media use in older adults provides an opportunity to reach patients who are eligible for cardiac rehabilitation but unable to attend traditional face-to-face groups. However, there is a paucity of research on cardiac patients' experiences and perspectives on using social media to support their health.

Objective: The aim of this study was to describe cardiac rehabilitation patients' experiences in using social media in general and their perspective on using social media, particularly Facebook, to support their cardiac health and secondary prevention efforts.

Methods: A mixed-methods study was undertaken among cardiac rehabilitation patients in both urban and rural areas. First, this study included a survey (n=284) on social media use and capability. Second, six focus group interviews were conducted with current Facebook users (n=18) to elucidate Facebook experience and perspectives.

Results: Social media use was low (28.0%, 79/282) but more common in participants who were under 70 years of age, employed, and had completed high school. Social media users accessed Web-based information on general health issues (65%, 51/79), medications (56%, 44/79), and heart health (43%, 34/79). Participants were motivated to invest time in using Facebook for "keeping in touch" with family and friends and to be informed by expert cardiac health professionals and fellow cardiac participants if given the opportunity. It appeared that participants who had a higher level of Facebook capability (understanding of features and the consequences of their use and efficiency in use) spent more time on Facebook and reported higher levels of "liking," commenting, or sharing posts. Furthermore, higher Facebook capability appeared to increase a participants' willingness to participate in a cardiac Facebook support group. More capable users were more receptive to the use of Facebook for cardiac rehabilitation and more likely to express interest in providing peer support. Recommended features for a cardiac rehabilitation Facebook group included a closed group, expert cardiac professional involvement, provision of cardiac health information, and ensuring trustworthiness of the group.

Conclusions: Cardiac health professionals have an opportunity to capitalize on cardiac patients' motivations and social media, mostly Facebook, as well as the capability for supporting cardiac rehabilitation and secondary prevention. Participants' favored purposeful time spent on Facebook and their cardiac health provides such a purpose for a Facebook intervention. The study results will inform the development of a Facebook intervention for secondary prevention of cardiovascular disease.

(J Med Internet Res 2017;19(9):e323) doi:10.2196/jmir.8081

KEYWORDS

cardiovascular disease; cardiac rehabilitation; qualitative research; secondary prevention; social media

Introduction

Coronary heart disease (CHD) is a leading cause of death worldwide [1]. A sizeable proportion of the burden of CHD is avoidable by targeting modifiable risk factors [2] such as tobacco use, high blood pressure, physical inactivity, poor diet, high body mass, and alcohol use [2]. Furthermore, attendance at group-based cardiac rehabilitation programs improves risk factors and subsequently reduces recurrent events [3]. Despite the benefits of cardiac rehabilitation, a vast majority of eligible patients, particularly those aged over 65 years, do not attend in-person services [4,5] because of geographical barriers, misconceptions, scheduling of programs, and conflicting demands [6,7].

Rapid growth in the use of the Internet and mobile phones, including smartphones, has emerged as a potential solution to improve access to cardiac rehabilitation. In 2016, 87% and 68% of adult populations from countries with advanced economies reported using the Internet and owning a smartphone, respectively [8]. These rates are even higher among the Australian adult population, with 93% using the Internet and 77% owning a smartphone [8]. Even among adults aged 65 years or older, the use of mobile phones and social media are becoming deeply embedded in everyday life [9]. Pervasive adoption of technology offers alternative opportunities for researchers and clinicians to engage with this hard-to-reach cardiac population [4,5] and establish services to meet the needs of future generations.

Platforms, such as Facebook, which enable social interactions through user-generated content, are collectively known as social media [10]. Social media usage among adults aged over 65 years has more than tripled from 2010 to 2015, from 11% to 35% in the United States [11]. Similarly, in Australia (2012-13), among the 46% of adults aged over 65 years who use the Internet, social media use is one of the four top Web-based activities [12]. As social influences are a primary factor in the adoption of health behaviors [13,14] and a core component of cardiac rehabilitation [15], adapting social support elements from cardiac rehabilitation programs for adjunct delivery via social media platforms is a logical progression to ensure flexible delivery and to maximize program reach [16].

The most commonly used platform for adults aged over 65 years in Australia [17] and the United States [9] is Facebook. As a result of the increase in Facebook popularity among this population, social support elements for electronic health (eHealth) or mobile health (mHealth) cardiac rehabilitation programs are often integrated into investigator-designed platforms as an adjunct to a multicomponent program [18,19]. However, these programs often do not detail the social media component or the role of user experience in their development. Understanding the user experience of Facebook and the potential to use this platform to achieve behavior change is essential to ensure the effectiveness of any program developed [20,21]. To date, however, there is only limited research on patient acceptance of existing social media platforms, such as Facebook, as a potential delivery modality for cardiac rehabilitation.

This study aims to describe cardiac rehabilitation patients' social media use; and among current users, their experience and perspectives on using Facebook to support their cardiac health through secondary prevention.

Methods

Design and Participants

This is a two phase mixed-methods study. Phase 1 comprised a survey of technology use in participants eligible for or currently attending cardiac rehabilitation. This survey was a component of a larger study (n=282) that screened cardiac patients regarding many aspects of their overall health technology use. Phase 2 comprised a series of focus groups with cardiac rehabilitation participants who were current Facebook users to elucidate their Facebook experience and the potential for using Facebook for the delivery and support of cardiac rehabilitation.

Phase 1 participants were cardiac patients recruited from cardiac hospital wards as inpatients or outpatients attending cardiac rehabilitation programs in New South Wales (NSW) at two metropolitan (7 hospitals) and three rural health districts (3 hospitals). Participants were candidates for the survey if they were eligible to be referred to cardiac rehabilitation and spoke and understood sufficient English for consent and questionnaire processes. Participants were excluded if they had a neurocognitive disorder or a major visual deficit.

Phase 2 participants were cardiac patients recruited from cardiac rehabilitation programs at one metropolitan and one rural hospital used in phase 1. Participants were eligible for this phase of the study, if in addition to the inclusion criteria for phase 1, they were currently using Facebook.

Ethical Considerations

The studies were approved by Northern Sydney Local Health District Human Research Ethics Committee in March 2016 (HREC ref: LNR/15/HAWKE/450), and all patients gave written informed consent before participation.

Data Collection

During phase 1, technology acceptance was assessed using a 20-item survey combining components of questionnaires developed by Edwards et al [22] and Illiger et al [23]. Questions included access to mobile technology, current usage patterns including features used, and confidence and interest in usage for health. Sociodemographic and clinical data were collected using a checklist developed for a previous study by the team [24]. This questionnaire was pilot-tested on 15 cardiac rehabilitation patients and modified to improve accuracy and specificity. Data collection for phase 1 took place from March to November 2016. This study reports only a subset of the data relating specifically to social media use in general.

During phase 2, the details of participants' Facebook experiences were collected using focus group interviews (n=6). The focus groups were aimed at understanding participants' Facebook experience, engagement, and confidence to identify factors associated with Facebook acceptance and to examine the potential for Facebook platforms to support their cardiac health and for secondary prevention options. A semistructured interview guide was used to elicit responses so that the question type and topics could vary according to participants' answers, and issues raised by previous groups could be added. Each of the focus groups began with introductions and an outline of processes and aims from the facilitator (SRP). Thereafter, participants were encouraged to express their opinions and were provided with sufficient time to do so. Participants were also asked to complete a short questionnaire that included sociodemographic and clinical characteristics, as well as questions on access to a broad range of social media platforms, current usage patterns, and confidence and interest in using social media for health purposes. An additional researcher (PG or SD) was present throughout the interview to assist with organization and audio-recording and to write field notes. The focus group recordings were transcribed verbatim for analysis.

Data Sources, Analysis, and Statistics

The survey data from phase 1 were analyzed using Statistical Package for the Social Sciences (SPSS) Statistics Version 22 (SPSS Inc). Descriptive statistics were used for continuous measures, including counts and percentages for participant demographic characteristics and technology use. Comparisons of demographic characteristics and technology use of non–social media users and social media users were conducted using chi-square tests.

The deidentified focus group transcripts were entered into NVivo version 10 software program (QSR International Pty Ltd). All focus groups were coded (SRP), allowing for data immersion and obtaining an overall sense of the data. An open coding approach was adopted, forming a general description of the research topic through generating codes and recurrent themes as they emerged. Discussion with additional researchers (AG and RG) familiar with the data finalized and confirmed emerging themes. Verbatim quotes illustrative of the themes and subthemes had been extracted. During this process of analyses, it became apparent that the participants' preferences for features of a Facebook group suitable for cardiac rehabilitation attendees were being revealed. Therefore, a secondary analysis of the transcripts and field notes was undertaken to identify these aspects and preferences for current and future Facebook groups by one researcher (PG). Indications (both positive and negative) of the strength of preference for different aspects of social media groups were assessed in transcripts and field notes and independently assessed by another researcher (RG). The aspects and the associated negative and positive strengths of preference were discussed between 3 researchers (PG, SRP, and RG) until consensus was reached. Symbols are used to indicate the overall strength of preference, with "−−−" indicating the features considered to be least desirable and "+++" indicating those features considered most desirable.

Results

Phase 1: Technology Survey

Participant Profile

Survey results indicated that 28% (79/282) of the cardiac rehabilitation participants in the sample used social media regularly (Table 1). In comparison with nonusers, social media users were younger; more than half (38/75; 51%) of social media users were aged ≤59 years compared with 17% (33/202) of nonusers. A higher proportion of social media users compared with non–social media users were in full-time employment (47% [37/79] vs 23% [47/203], *P*<.001), had a secondary education qualification or higher (76% [60/79] vs 58% [118/203], *P*=.04), and regularly used smartphones (90% [71/79] vs 53% [108/203], *P*<.001) and tablets (64% [50/79] vs 34% [68/203], *P*<.001). Almost two-thirds of social media users compared with one-third of non–social media users were accessing information on health conditions (*P*<.001). Social media users compared with non–social media users were significantly more likely to use technology to access information on medications, heart conditions, heart treatments, and lifestyle changes. Additionally, 30% (24/79) of social media users compared with 12% (24/203) of nonusers were willing to communicate with health professionals using technology (*P*<.001). There was no statistically significant difference between social media users and nonusers with their willingness to use technology to communicate with other cardiac participants (8% [6/79] vs 4% [8/203], *P*=.17).

Phase 2: Focus Groups

Participants' Profile

Of the 20 participants invited to participate in the focus groups, 18 took part (72% urban). Focus group participants had a mean age of 63 years (standard deviation [SD] 12; range: 38-79 years). The majority were males (13/18), had some form of tertiary education (15/18; mean years of education=14.4, SD=2.7), and were white (13/18). The 5 non-white participants were of Indian, European, Aboriginal, or Pacific Islander descent. The majority of the participants were working part-time (4/18) or were retired (9/18) and were living with a partner (12/18). Among this group, one-third each had been admitted for myocardial infarctions (6/18), or percutaneous coronary interventions (6/18), or coronary artery bypass grafting or valve replacement surgery (5/18).

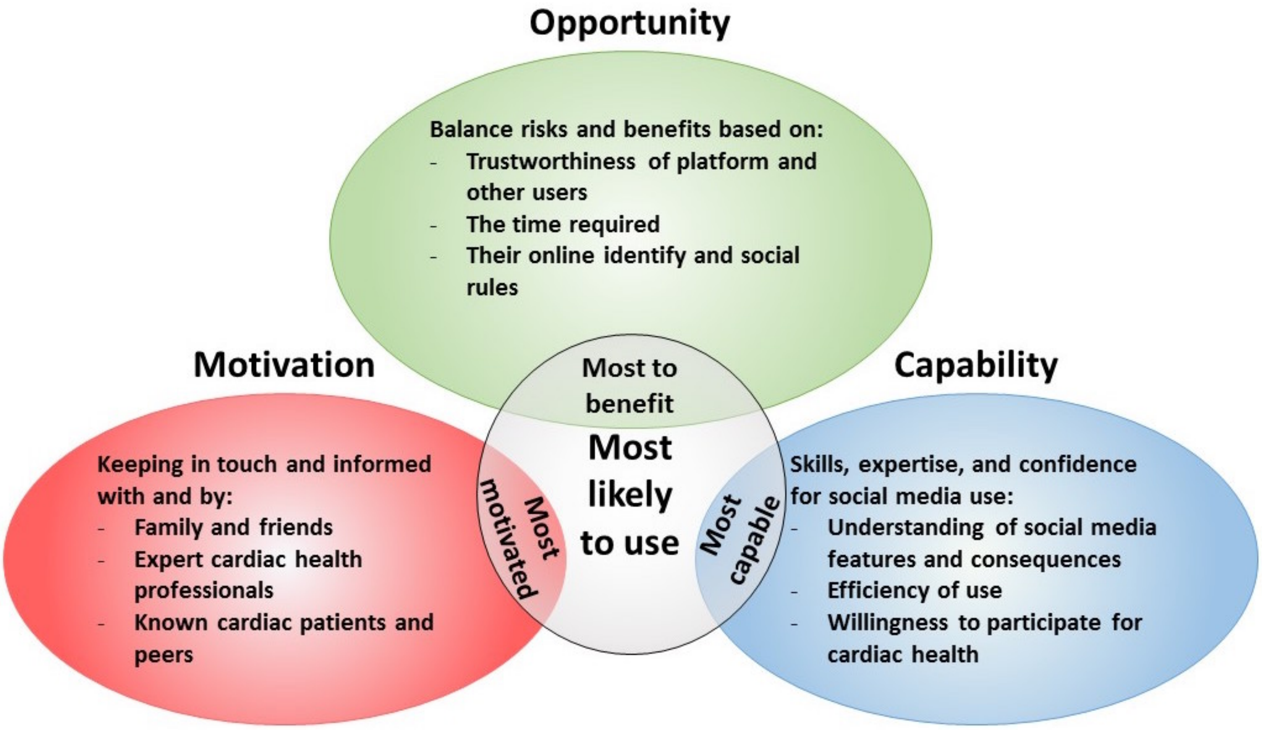
Table 1. Phase 1 survey: demographics and technology use by cardiac rehabilitation participants compared for social medial use (N=282).

Demographics and technology use	Non–social media users (N=203) n (%)	Social media users (N=79) n (%)	*P* value
Demographic characteristics			
Age in years [a], n (%)			
49 or younger	6 (3.0)	11 (14)	<.001
50-59	27 (13.4)	27 (36)	
60-69	70 (34.7)	22 (29)	
70-79	73 (36.1)	11 (14)	
80 or older	26 (12.9)	4 (5)	
Gender, n (%)			
Female	57 (28.1)	21 (27)	.80
Language spoken at home, n (%)			
English	185 (91.1)	73 (92)	.55
Education level, n (%)			
≥year 12[b]	118 (58.1)	60 (76)	.004
Location, n (%)			
Metropolitan	144 (70.9)	69 (87)	.18
Employed, n (%)			
Full-time or part-time	47 (23.2)	37 (47)	<.001
Technology use			
Device, n (%)			
Tablet or iPad	68 (33.5)	50 (63)	<.001
Smartphone	108 (53.2)	71 (90)	<.001
Accessing information, n (%)			
Health conditions	66 (32.5)	51 (65)	<.001
Medications	54 (26.6)	44 (56)	<.001
Heart condition	36 (17.7)	34 (43)	<.001
Heart treatment	32 (15.8)	29 (37)	<.001
Lifestyle changes	34 (16.7)	30 (38)	<.001
Communication, n (%)			
Health professionals	24 (11.8)	24 (30)	<.001
Other cardiac patients	8 (3.9)	6 (8)	.17

[a] One non–social media user did not report age, and 4 social media users did not report age.

[b] Secondary education or higher.

Figure 1. Social media use themes for cardiac rehabilitation patients.



The short questionnaire completed by participants before the start of the focus group found that the most commonly used social media site was Facebook (16/18), followed by Instagram (7/18), LinkedIn (5/18), and Twitter (5/18). The average number of social media sites per participants used was three (range: 1-5). The majority were multidevice users, with 15 of 18 using a computer or laptop, tablet, and a smartphone. Most participants accessed social media sites on their smartphone (13/18) and in combination with one other device—their computer or laptop (11/18) or their tablet (13/18).

Focus Group Themes

Participant discussion of their current Facebook use appeared to involve the interaction of three components, namely, motivation, opportunity, and apparent capability (Figure 1). As a result, the participants who were most likely to use Facebook demonstrated high levels of all three key attributes portrayed in Figure 1, that is, those most likely to use Facebook were also the most motivated, the most capable, and identified the most benefits from the potential opportunities Facebook provides. Participants were motivated to keep in touch with known family and friends, and they recognized that Facebook provided an opportunity for such social interactions. Participants' use of Facebook strongly related to their apparent capability to use Facebook and their perceived benefits of social media use. For example, those perceiving fewer benefits of Facebook appeared or judged themselves to have lower Facebook capability and less motivation to use Facebook to keep in touch (ie, used other forms of communication). If given the opportunity to use a cardiac Facebook group, participants indicated they were motivated to do so to keep informed by expert health professionals, but again their willingness to participate was influenced by their Facebook capability.

Motivation

Participants described their main motivation for using Facebook generally was to "keep in touch" with family and friends, being informed, and informing each other of current life events. Facebook provided support for existing relationships as opposed to developing new relationships. For example:

> *I use it to keep in touch with friends, and groups, and uh, it's very good for keeping touch with family that are living out interstate...It's good. It helps you to keep in touch with people, finding out what they all know and what they're doing.* [Male, 69, focus group 1]
>
> *The thing I like the most is it keeps me in contact, it lets me see what other people are doing which I normally wouldn't do. Like, I've got lots of extended family and cousins and different things and so they'll put things on about their kids or somebody's gone on holidays and all that and I think that's good because we're all so busy nobody has time to make phone calls these days.* [Female, 59, focus group 2]

A secondary motivation for Facebook use described by younger participants as well as those who were older but "early adopters" was that it helped to fill in spare time. One of the participants stated:

> *The beauty of the phone is you can, you know, if you're waiting for a train or waiting for anything you can entertain yourself as much or as little as you like.* [Male, 69, focus group 1]

However, the motivation to keep in touch with family and friends seemed to be the most powerful, influencing their development of skills to engage and the way they engaged with Facebook.

Opportunity

From the focus groups, it appeared that participants harness the opportunity Facebook provided to keep in touch with family and friends in different ways and carefully balanced risks and benefits in an ongoing way. For instance, participants weighed the opportunity to keep informed of family events against the risk (as they perceive it) of excessive advertisements. Similarly, the balance of the opportunity to engage in a conversation with family and friends was weighed against the time required to filter and make posts and the risk of generating unwelcome responses. Additionally, participants weighed up these stated benefits with the time required to develop and define their online identity and to understand the rules of social engagement on Facebook.

Most participants mentioned that their use of Facebook involved observing and sharing personal photos, "liking" photos or updates of family and friends, and communicating through private messaging features but seldom communicated through groups, as stated below:

> *I mostly browse but if it's something that I really want to see, like I've got friends that live overseas and things like that and they'll post cute photos of their kids and stuff, I definitely will like that and read it or something. Yeah, it's a way of letting people know that you know what's happening.* [Female, 59, focus group 2]
>
> *I use Facebook a lot and Facebook Messenger, which I like because it's a private conversation and you don't have all the other things that you have to scroll through.* [Female, 79, focus group 1]

Participants had varying levels of trust of both the Facebook platform and of other users. Trust was informed by other, non-Facebook (eg, email and Internet banking) Web-based experiences removed from Facebook platforms and impacted the way participants navigated and used the platform. Frequent, confident Facebook users reported developing strategies to deal with trust issues while still using the platform, which is echoed by the following:

> *...even though I'm on Facebook and Wi-Fi a lot, every single day, I would not share anything that I consider to be personal, private, financial, medical information through that media and never have. Because you can't guarantee that it's secure...but I've got to be very very sure that I'm using something that has a secure transaction.* [Female, 58, focus group 3]
>
> *I take it with a grain of salt. Facebook is one of those mediums where people can write whatever they want, there's no factual evidence that supports what their saying, it's just their opinion and I take it as that, as their opinion I never take it as complete fact.* [Male, 38, focus group 3]
>
> *Sometimes there's too much on there and you just think "oh well, whatever I've missed I've missed I don't have time." But yeah I find...there [is] a lot other stuff on there that you don't want to see.* [Female, 59, focus group 2]

However, less frequent or confident Facebook users did not describe having strategies in place to deal with their trust issues and reported less control of their online personal space. This increased their time on Facebook and appeared to negatively affect their Facebook experience, as expressed by some of the participants:

> *What I don't like about Facebook is how you can say what you want, you can swear, every word known to man, now I'm not a prude and they get away with it; if you put that in the paper you'd have, you'd be sued. How do they allow that? That's my question...I think it's putrid, it's filth. It's putrid. But it's technology.* [Male, 73, focus group 6]
>
> *It's like 4 people come up you don't know from a bar of soap wanting to be your friends, what's all that?...I don't know who they are.* [Male, 59, focus group 6]

Participants expressed the view that spending a large amount of time on Facebook was undesirable because of the potential negative perceptions by others. Increased time spent on Facebook is perceived to signal that the person has no other interests or they are oversharing personal information or personal opinions of current popular culture or current affairs, irrespective of trust and capability; this was reiterated by some of the participants:

> *I just think a person who's posting a lot of stuff is spending a lot of time on Facebook...I think [they] should be doing something more with [their] life. That's my personal opinion.* [Female, 79, focus group 1]
>
> *I use it a lot I guess, but I don't like to sit on it all the time, I like to get on and off and just enjoy what I am doing and then if I am going to do something then I'll go on it, because I can't just stand sitting on the computer, there's got to be a reason for being there.* [Female, 58, focus group 5]

However, it appears that participants would reach a point or threshold when they considered they had invested too much time, particularly when using handheld devices, and they then described developing strategies to reduce use. For example:

> *Actually, I just deleted [Facebook] on my phone because I got sick of...I'd look at it too much. So I decided I'm only going to look at it on my iPad, which is not with me all the time.* [Female, 62, focus group 1]

Participants' online Facebook behaviors appear to be a reflection of the rules of behavior they consider acceptable in their social group and in relation to communication with their friends and family. This included developing strategies to tolerate specific family and friends and not changing their reactions for an online context, as noted by some participants:

> *You can unfollow and still be friends, I do that a lot. If someone keeps posting constantly you just click unfollow and you stay friends. They don't know that you're not seeing it.* [Female, 62, focus group 1]

> *I wouldn’t say something on Facebook to anybody that I wouldn’t say to their face.* [Male, 63, focus group 6]

Capability

Participants’ engagement and perception of Facebook were influenced by their apparent capability. Comments by participants revealed that their capability of understanding and using Facebook fell on a continuum of skill, expertise, and confidence, which then influenced whether they used the opportunity Facebook provided.

Participants who were apparently less capable with regard to Facebook use revealed their lower level of understanding of Facebook features in distinct ways in their comments during the focus groups, such as being less able to control the number and type of posts and notifications they received. As a consequence, they were more likely to be distracted and annoyed by common Facebook features and the amount of perceived trivial communication. This also increased the amount of time they spent on Facebook to achieve what they wanted, thus perceiving it as a “necessary evil” (Male, 73, focus group 6). Oftentimes, less capable Facebook users were unable to distinguish between Facebook platforms and the wider Internet. Their sense of safety or risk was then falsely enhanced, extending beyond what is possible. For example, one of the participants noted:

> *I just got loaded with stuff and a lot of it was interesting, like I like bushwalking and I put down my interests. But then there was just so much going through that, I was bombarded, but I kept seeing things like “yes I’d like to do that one day” but in the end I just had to delete it, it was just overwhelming.* [Female, 79, focus group 1]

Participants who were apparently more capable revealed a better understanding of Facebook features and the consequences of their use through their easy navigation of content with limited distractions. They also described their communication through the confident use of multiple features:

> *If I’m just putting a post up myself or answering one of their posts then that would in the public domain and then when it came around to if we needed to discuss things of a family nature then that would go to private messaging.* [Male, 56, focus group 4]

> *If I’ve got something sensible to say I do, but a lot of people make silly comments and unnecessary comments and it is just time consuming.* [Male, 69, focus group 1]

Differences in participants’ apparent capability were also revealed through the language used to describe Facebook actions and the desire for simplicity. Some participants were aware of their lower level of capability in using Facebook and discussed that their lower skill level was a result of learning Facebook at an older age:

> *I mean, it’s [got to] be simple for our age group, I think. Like, that we don’t have to have other people to do it. You know, it’s [got to] be easy enough.* [Female, 62, focus group 1]

> *A lot of it is coming through the young people, the young ones with their work. They get very involved and learn about all these things, and we as the older generation learn from them. But it’s not something that we learned when we were at work so we’re kind of on the edge and catching up with the young ones. That’s what I find I do at times.* [Female, 58, focus group 5]

Potential Use of Facebook for Cardiac Rehabilitation

No participant raised the potential for using Facebook for their own health or specifically for cardiac health; however, after being prompted to consider this possibility, a lot of discussion ensued. Most participants discussed that Facebook, in general, would provide an opportunity for continued interaction with known expert cardiac health professionals and cardiac rehabilitation peers to receive further cardiac specific information and be provided with personalized cardiac information. Participants who were more capable Facebook users recognized the benefits of Facebook private groups for this purpose and were more willing to participate in a Facebook group and to help others through peer-to-peer learning, which is echoed by the following:

> *You know it is information at hand when you need it. You don’t have to physically go into a place and have a chat with someone, it’s there. Something that you use often and it’s available at a swipe. Someone can be there to answer any questions that you’ve got. I think it’s a great idea.* [Male, 38, focus group 3]

> *The beauty of it is it’s a very quick way of getting out information and it costs nothing, and it’s just very effective.* [Male, 69, focus group 1]

> *The other thing that may be a good idea either for myself or others is to be available to a group like that to mentor...because there are things that I’ve gone through, there are things that each of you, I'm sure had things that have been really beneficial, things that have been a struggle—to be a mentor to somebody, or a buddy, I don’t know call it what you will, that’s something I would be happy to do.* [Female, 58, focus group 3]

Participants with lower perceived benefits of Facebook and apparently lower capabilities were still willing to “*give it a go*” (Male, 79, focus group 5). Another participant noted:

> *I think that’s why people are cynical about Facebook...but if it is dealing with things that are, issues, whatever, then that’s a good thing. Like cardiac rehabilitation or something like that.* [Male, 63, focus group 6]

Some of the participants expressed their hesitation about the level of sharing of personal health information required and privacy and confidentiality of shared health information:

> *Yeah, I think it would do...be very good because you can share. Once you start sharing the tension and the*

> *unknown is not as great and that releases any stress that you would generally hold on and wonder what's happening.* [Female, 69, focus group 3]
>
> *As long as there's privacy and I could trust it and it could be guaranteed then yeah definitely. So there's a trust issue there. So it couldn't be accessed by people flogging products or services. My dealing is with you and the hospital and not with the mob that's selling water bottles for example.* [Male, 63, focus group 1]

Recommendations of Features When Creating a Facebook Intervention for Cardiac Rehabilitation

In light of both the findings described above and the secondary analyses of preferences for Facebook, a series of recommendations have been generated regarding features to consider when creating a cardiac rehabilitation Facebook group (Table 2**)**. Three categories need to be taken into consideration in relation to platform settings, the role of the moderator, and group interaction. Participants strongly preferred a closed or secret group, and they wanted to be able to access the platform on multiple devices. Strong preferences were also indicated for having an expert cardiac professional act as moderator and that content should focus on cardiac health information and provide link-outs to relevant material. Participants also emphasized the importance of trustworthiness of the group, whereas excessive notifications, advertisements, and hostile or augmentative posts were not favored.

Table 2. Facebook group preferences.

Categories and features	Preference[a]
Platform settings	
"Closed" or "Secret" group	+++
Multiple device access	++
Presence of advertisements	−
Excessive notifications	−−
Moderator role	
Expert cardiac professional	+++
Cardiac health information	++
Cardiac information link-outs	++
Group interaction	
Trustworthiness of group	++
Ability to post or comment	+
Irrelevant posts	−
Hostile or argumentative posts	−−

[a] +++ Most desirable; −−− Least desirable.

Discussion

Principal Findings

This mixed-methods study provides strong indications that a cardiac Facebook group has potential for cardiac rehabilitation purposes and provides a set of recommended features to consider when designing such a Facebook group. To our knowledge, this is the first study to elucidate cardiac rehabilitation participants' perceptions of and engagement with social media and the potential for using Facebook for their cardiac health. Social media use was low but more common in participants who were under 70 years of age, employed, and had completed high school. It appeared cardiac rehabilitation participants' Facebook engagement is influenced by a combination of personal motivation, opportunity, and capability of using the platform. When applied to cardiac rehabilitation, participants could be motivated to use Facebook because of the opportunity it provided to "keep in touch" with a known expert cardiac health professional and to further improve their cardiac health. More capable Facebook users may also be open to providing peer support in a social media group by acting as champions. The findings of this study will be used to develop a Facebook intervention for persons eligible for cardiac rehabilitation.

Key prerequisites in social media use in older adults were identified previously in a review by Leist [16] and include adoption of information and communications technology and social media–related knowledge, attitudes toward social media, and motivations for social media use. All participants in this study demonstrated varying levels of these prerequisites. A combination of their motivation to use Facebook for supporting their existing relationships and their capability of Facebook use seems to influence the participants' Facebook-related knowledge and attitudes toward Facebook. Also similar to Leist [16], we found barriers such as lack of capability and skeptical attitudes that were described as more constraining with increasing age [25]. In this study, participants aged 70 years or older were much less likely to be using social media in general, which might limit the potential use of Facebook for cardiac rehabilitation. For instance, given that the mean age of heart attack participants in the study context of Australia is 67 years (SD 15), it is likely that many participants will not be capable of using a cardiac Facebook group option [26]. However, it is highly likely that age-related barriers to social media use will diminish rapidly among future generations of cardiovascular disease (CVD) participants because of the rapid rise and increased use of smartphone and social media use among younger generations and the influence of younger family members on older people [8]. Moreover, encouraging older people to use social media and emphasizing the benefits of social media use would likely result in more engagement with and positive attitudes toward social media [27]. Older adults are still likely to use the Internet to seek health information and to connect to individuals experiencing similar conditions, mainly through discussion groups [11].

Social media–based discussion groups present an important new resource for health researchers to understand, given the primary role of social networks in the adoption of health behaviors [13,14]. Social media features such as information sharing allow social reinforcement to occur in real time [28,29]. However, the aforementioned barriers to social media use identified in this study may prevent online discussion groups from achieving peer-to-peer interactions, which are comparable to traditional in-person support groups. This may be overcome by supporting less capable users through peer-led mentoring, which is recognized as a potentially effective resource; however, it remains underinvestigated in chronic disease management [30]. Participants in this study, who were capable Facebook users, volunteered to act as mentors to less capable users to become more acquainted with Facebook. This approach was also suggested by Leist [16], who identified Internet-savvy members of online communities who were willing to act as mentors and at times, moderators. Such an approach has additional benefits, as it may promote continuous engagement of users in an online group, which is likely to maximize benefits [16].

Our study findings indicate that some participants are very capable in terms of engaging with personal interest groups via Facebook. However, none of the participants in this study recognized the potential Facebook provided for continued interaction with known cardiac experts and peers until prompted. Previous studies have reported that lack of personal relevance to a social media platform resulted in negative attitudes and lower engagement [20]. The use of well-known and existing social media platforms such as Facebook for interventions for less capable persons may be important for continued engagement [16]. Furthermore, the use of personally relevant social media platforms may have the potential to increase engagement in capable users, as found in this study—participants' favored purposeful time spent on Facebook.

The majority of studies to date reporting social media interventions for health-related behavior change have documented outcomes rather than the process of intervention development. In 2014, a systematic review of studies, which incorporated Web-based social media platforms, reported that there is modest evidence that behavior change interventions targeting key modifiable health behaviors may be effective, but this review noted that this field of research is still in its infancy [21]. This review highlighted the importance of maximizing retention and engagement within social networking interventions and the limited information reported on the long-term effectiveness of the resultant behavior change. Of the 10 interventions included in the review [21], only five used an existing social media platform [31-35], with one using Twitter [34] and four using the private group feature of Facebook [31-33,35]. These four studies utilized private Facebook groups as an adjunct to an existing intervention to enhance social support for participants. None reported prior formative research on the social media experience and perspectives of their target group, and none of these studies involved participants with CVD.

The effectiveness of social media use as a delivery medium for the secondary prevention of cardiac rehabilitation remains unclear. Two systematic reviews investigating the effectiveness of mHealth interventions for the secondary prevention of CVD are available but are limited in that none of the included studies have investigator-built online social support elements or social networking via established social media platforms [15,36]. One digital cardiac rehabilitation intervention, designed as an adjunct to traditional in-person services, has been tested [37]. This intervention contained a Web-based application as well as a mobile phone app for personalized health assistance, coupled with a social reinforcement network that aimed to encourage the adoption and maintenance of a healthier lifestyle. However, no process evaluation data are available to understand how participants interacted with the social reinforcement network or the potential influence on outcomes. Furthermore, a study protocol describing a remotely delivered exercise-based cardiac rehabilitation mHealth program with a secure social support component has been designed and incorporates self-efficacy elements; however, it has not yet been implemented [38].

Insights of current Facebook use by participants attending cardiac rehabilitation in this study have provided an essential foundation for the design of a Facebook intervention for the secondary prevention of CVD. mHealth interventions, including those with social media elements, underpinned by behavioral theories have been shown to be more effective compared with mHealth interventions with no theoretical underpinning [39,40]. The themes from this study, namely, capability, opportunity, and motivation and interaction between them, have commonality with both social cognitive theory and self-efficacy [41] and with the behavior change wheel [42]. On the basis of evidence from interventions that are not based on social media, integrating principles from these behavior change theories for the secondary prevention of CVD may significantly increase the likelihood of success and therefore has potential for social media–based interventions. In addition, frameworks for adapting existing social media delivery will be taken into consideration [43], as well as the process of communicating health information via social media [44,45].

Strengths and Limitations

The strength of this formative research is the use of both quantitative and qualitative data sources. A relatively diverse sample of both metropolitan and rural participants was represented in both phases of the study. However, it should be emphasized that the nature of the focus group study and qualitative research limits the generalizability of the results to the wider cardiac rehabilitation population, including those patients who are eligible for cardiac rehabilitation yet do not attend as well as the rural patients. It is possible that these subgroups could offer useful insights into accessibility of services; however, they are difficult populations to recruit and engage with.

Conclusions

In this study, we have shown that group-based social media programs, using existing platforms such as Facebook, may offer an opportunity to access and engage cardiac rehabilitation patients. One in 4 cardiac rehabilitation patients are current users of Facebook—a proportion only likely to increase in the future, given the rapidly growing use of social media in older adults. Considering that cardiac rehabilitation patients who use social media are already frequently accessing information on their health, cardiac conditions, and lifestyle, it is important to leverage this potential with appropriate theory-based support. Capitalizing on the social media potential for cardiac patient support depends on understanding patients' motivation to keep in touch with known health professionals and peers, efficiency of use **,** and ensuring that the benefits outweigh potential risks. This study has brought greater understanding of Facebook use and the perceptions of and engagement with Facebook by adults with CVD. As such, we have generated a set of recommendations for consideration when establishing cardiac rehabilitation Facebook groups. Future work will incorporate the findings and recommendations to develop a Facebook intervention to support adults with CVD to improve their modifiable risk factors and to lower their chances of further events.

Acknowledgments

Phase 1 of this research was funded by the NSW Nursing and Midwifery Office Innovations Scholarship and seeding funds from the NSW Heart Foundation. Phase 2 of this research was funded by a NSW Cardiovascular Research Network collaborative project grant.

The authors wish to thank Royal North Shore, Ryde, Manly, Hornsby Ku-Ring-Gai, Prince of Wales, St. George, John Hunter and Wagga Wagga Hospitals and Coffs Harbour Health Campus as well as the staff and patients of the cardiac rehabilitation programs at these institutions for their contributions to the questionnaire component of this study. In particular, the authors wish to thank Julie Belshaw, Ann Kirkness, Jenny Fildes, Dawn McIvor, Joanne Leonard, and Sheryl Bowen for their assistance with recruitment for this component of this study.

Authors' Contributions

SRP, AG, BF, BO, LN, SD, GP, DD, AB, PP, and RG designed the questionnaire and focus group questions used in phase 2 of this study. SRP moderated the focus groups. PG and SD took field notes. PG transcribed the focus groups. SRP, AG, PG, and RG analyzed the data. KR, LS, HG, and RG designed the questionnaire used in phase 1 of this study. SRP drafted the manuscript with input from all coauthors. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

References

1. Australian Institute of Health and Welfare. Australian burden of disease study: impact and causes of illness and death in Australia 2011. In: Australian Burden of Disease Study series no. 3. BOD 4. Canberra: AIHW; 2011.
2. Unal B, Critchley JA, Capewell S. Explaining the decline in coronary heart disease mortality in England and Wales between 1981 and 2000. Circulation 2004 Mar 09;109(9):1101-1107 [FREE Full text] [doi: 10.1161/01.CIR.0000118498.35499.B2] [Medline: 14993137]
3. National Heart Foundation of Australia. Secondary prevention of cardiovascular disease. Canberra: National Heart Foundation of Australia; 2010.
4. Cortés O, Arthur HM. Determinants of referral to cardiac rehabilitation programs in patients with coronary artery disease: a systematic review. Am Heart J 2006 Feb;151(2):249-256. [doi: 10.1016/j.ahj.2005.03.034] [Medline: 16442885]
5. Daly J, Sindone AP, Thompson DR, Hancock K, Chang E, Davidson P. Barriers to participation in and adherence to cardiac rehabilitation programs: a critical literature review. Prog Cardiovasc Nurs 2002;17(1):8-17. [Medline: 11872976]
6. Clark AM, King-Shier KM, Thompson DR, Spaling MA, Duncan AS, Stone JA, et al. A qualitative systematic review of influences on attendance at cardiac rehabilitation programs after referral. Am Heart J 2012 Dec;164(6):835-45.e2. [doi: 10.1016/j.ahj.2012.08.020] [Medline: 23194483]
7. Neubeck L, Freedman SB, Clark AM, Briffa T, Bauman A, Redfern J. Participating in cardiac rehabilitation: a systematic review and meta-synthesis of qualitative data. Eur J Prev Cardiol 2012 Jun;19(3):494-503. [Medline: 22779092]

8. Pew Research Center. 2016. Smartphone ownership and internet usage continues to climb in emerging economies URL: http://www.pewglobal.org/2016/02/22/smartphone-ownership-and-internet-usage-continues-to-climb-in-emerging-economies/ [accessed 2016-03-15] [WebCite Cache ID 6q9XJ1jiy]
9. Pew Research Center. 2016. Social media update 2016 URL: http://www.pewinternet.org/2016/11/11/social-media-update-2016/ [accessed 2017-01-30] [WebCite Cache ID 6q9XCJ4kT]
10. Safko L. The Social Media Bible: Tactics, Tools, and Strategies for Business Success. Hoboken, NJ: Wiley; 2012.
11. Perrin A. Pew Research Center. 2015. Social media usage: 2005-2015 URL: http://www.pewinternet.org/2015/10/08/social-networking-usage-2005-2015/ [accessed 2016-08-09] [WebCite Cache ID 6q9XMooTa]
12. Australian Bureau of Statistics. ABS. 2016. Older persons internet use URL: http://www.abs.gov.au/ausstats/abs@.nsf/Lookup/B10BDF0266D26389CA257C89000E3FD8?opendocument [accessed 2017-05-02] [WebCite Cache ID 6q9XQWswq]
13. Pampel FC, Krueger PM, Denney JT. Socioeconomic disparities in health behaviors. Annu Rev Sociol 2010 Aug;36:349-370 [FREE Full text] [doi: 10.1146/annurev.soc.012809.102529] [Medline: 21909182]
14. Smith KP, Christakis NA. Social networks and health. Annu Rev Sociol 2008 Aug;34(1):405-429. [doi: 10.1146/annurev.soc.34.040507.134601]
15. Beatty AL, Fukuoka Y, Whooley MA. Using mobile technology for cardiac rehabilitation: a review and framework for development and evaluation. J Am Heart Assoc 2013 Nov 01;2(6):e000568. [doi: 10.1161/jaha.113.000568]
16. Leist AK. Social media use of older adults: a mini-review. Gerontology 2013;59(4):378-384 [FREE Full text] [doi: 10.1159/000346818] [Medline: 23594915]
17. Sensis. Sensis Social Media Report 2016: How Australian people and businesses are using social media. Melbourne, Australia: Sensis; 2016.
18. Neubeck L, Lowres N, Benjamin EJ, Freedman SB, Coorey G, Redfern J. The mobile revolution--using smartphone apps to prevent cardiovascular disease. Nat Rev Cardiol 2015 Jun;12(6):350-360. [doi: 10.1038/nrcardio.2015.34] [Medline: 25801714]
19. Neubeck L, Redfern J, Fernandez R, Briffa T, Bauman A, Freedman SB. Telehealth interventions for the secondary prevention of coronary heart disease: a systematic review. Eur J Cardiovasc Prev Rehabil 2009 Jun;16(3):281-289. [doi: 10.1097/HJR.0b013e32832a4e7a] [Medline: 19407659]
20. Xie B, Watkins I, Golbeck J, Huang M. Understanding and changing older adults' perceptions and learning of social media. Educ Gerontol 2012 Apr 01;38(4):282-296 [FREE Full text] [doi: 10.1080/03601277.2010.544580] [Medline: 22639483]
21. Maher CA, Lewis LK, Ferrar K, Marshall S, De BI, Vandelanotte C. Are health behavior change interventions that use online social networks effective? a systematic review. J Med Internet Res 2014;16(2):e40 [FREE Full text] [doi: 10.2196/jmir.2952] [Medline: 24550083]
22. Edwards L, Thomas C, Gregory A, Yardley L, O'Cathain A, Montgomery AA, et al. Are people with chronic diseases interested in using telehealth? a cross-sectional postal survey. J Med Internet Res 2014;16(5):e123 [FREE Full text] [doi: 10.2196/jmir.3257] [Medline: 24811914]
23. Illiger K, Hupka M, von JU, Wichelhaus D, Albrecht U. Mobile technologies: expectancy, usage, and acceptance of clinical staff and patients at a university medical center. JMIR Mhealth Uhealth 2014;2(4):e42 [FREE Full text] [doi: 10.2196/mhealth.3799] [Medline: 25338094]
24. Gallagher R, Belshaw J, Kirkness A, Warrington D, Sadler L, Roach K. Evaluation of a brief educational intervention to improve knowledge of sublingual nitroglycerin in cardiac rehabilitation patients. Eur J Cardiovasc Nurs 2013 Dec;12(6):529-535. [doi: 10.1177/1474515112473694] [Medline: 23315128]
25. Lee B, Chen Y, Hewitt L. Age differences in constraints encountered by seniors in their use of computers and the internet. Comput Human Behav 2011 May;27(3):1231-1237. [doi: 10.1016/j.chb.2011.01.003]
26. Chew DP, French J, Briffa TG, Hammett CJ, Ellis CJ, Ranasinghe I, et al. Acute coronary syndrome care across Australia and New Zealand: the SNAPSHOT ACS study. Med J Aust 2013 Aug 05;199(3):185-191. [Medline: 23909541]
27. Lehtinen V, Näsänen J, Sarvas R. “A Little Silly and Empty-Headed” - Older Adults' understandings of social networking sites. 2009 Presented at: Proceedings of the 23rd British HCI Group Annual Conference on People and Computers: Celebrating People and Technology Cambridge, British Computer Society; September 01-05, 2009; Cambridge, United Kingdom.
28. Bennett GG, Glasgow RE. The delivery of public health interventions via the internet: actualizing their potential. Annu Rev Public Health 2009;30:273-292. [doi: 10.1146/annurev.publhealth.031308.100235] [Medline: 19296777]
29. Buis LR. The potential for web-based social network sites and self-regulation for health promotion. Am J Health Promot 2011;26(2):73-76. [doi: 10.4278/ajhp.110223-CIT-83] [Medline: 22040387]
30. Funnell MM. Peer-based behavioural strategies to improve chronic disease self-management and clinical outcomes: evidence, logistics, evaluation considerations and needs for future research. Fam Pract 2010 Jun;27(Suppl 1):i17-i22 [FREE Full text] [doi: 10.1093/fampra/cmp027] [Medline: 19509083]
31. Cavallo DN, Tate DF, Ries AV, Brown JD, DeVellis RF, Ammerman AS. A social media-based physical activity intervention: a randomized controlled trial. Am J Prev Med 2012 Nov;43(5):527-532 [FREE Full text] [doi: 10.1016/j.amepre.2012.07.019] [Medline: 23079176]

32. Foster D, Linehan C, Kirman B, Lawson S, James G. Motivating physical activity at work: using persuasive social media for competitive step counting. 2010 Oct 06 Presented at: Proceedings of the 14th International Academic MindTrek Conferencenvisioning Future Media Environments; October 06-08, 2010; Tampere, Finland.
33. Napolitano MA, Hayes S, Bennett GG, Ives AK, Foster GD. Using Facebook and text messaging to deliver a weight loss program to college students. Obesity (Silver Spring) 2013 Jan;21(1):25-31. [doi: 10.1002/oby.20232] [Medline: 23505165]
34. Turner-McGrievy G, Tate D. Tweets, apps, and pods: results of the 6-month mobile pounds off digitally (Mobile POD) randomized weight-loss intervention among adults. J Med Internet Res 2011;13(4):e120 [FREE Full text] [doi: 10.2196/jmir.1841] [Medline: 22186428]
35. Valle CG, Tate DF, Mayer DK, Allicock M, Cai J. A randomized trial of a Facebook-based physical activity intervention for young adult cancer survivors. J Cancer Surviv 2013 Sep;7(3):355-368 [FREE Full text] [doi: 10.1007/s11764-013-0279-5] [Medline: 23532799]
36. Park LG, Beatty A, Stafford Z, Whooley MA. Mobile phone interventions for the secondary prevention of cardiovascular disease. Prog Cardiovasc Dis 2016;58(6):639-650. [doi: 10.1016/j.pcad.2016.03.002] [Medline: 27001245]
37. Widmer RJ, Allison TG, Lerman LO, Lerman A. Digital health intervention as an adjunct to cardiac rehabilitation reduces cardiovascular risk factors and rehospitalizations. J Cardiovasc Transl Res 2015 Jul;8(5):283-292 [FREE Full text] [doi: 10.1007/s12265-015-9629-1] [Medline: 25946990]
38. Rawstorn JC, Gant N, Meads A, Warren I, Maddison R. Remotely delivered exercise-based cardiac rehabilitation: design and content development of a novel mhealth platform. JMIR Mhealth Uhealth 2016;4(2):e57 [FREE Full text] [doi: 10.2196/mhealth.5501] [Medline: 27342791]
39. Glanz K, Bishop DB. The role of behavioral science theory in development and implementation of public health interventions. Annu Rev Public Health 2010;31:399-418. [doi: 10.1146/annurev.publhealth.012809.103604] [Medline: 20070207]
40. Webb TL, Joseph J, Yardley L, Michie S. Using the internet to promote health behavior change: a systematic review and meta-analysis of the impact of theoretical basis, use of behavior change techniques, and mode of delivery on efficacy. J Med Internet Res 2010;12(1):e4 [FREE Full text] [doi: 10.2196/jmir.1376] [Medline: 20164043]
41. Bandura A. Self-efficacy: toward a unifying theory of behavioral change. Psychol Rev 1977 Mar;84(2):191-215. [Medline: 847061]
42. Michie S, van SM, West R. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. Implement Sci 2011;6:42 [FREE Full text] [doi: 10.1186/1748-5908-6-42] [Medline: 21513547]
43. Pagoto S, Waring ME, May CN, Ding EY, Kunz WH, Hayes R, et al. Adapting behavioral interventions for social media delivery. J Med Internet Res 2016;18(1):e24 [FREE Full text] [doi: 10.2196/jmir.5086] [Medline: 26825969]
44. Kite J, Foley BC, Grunseit AC, Freeman B. Please like me: facebook and public health communication. PLoS One 2016 Sep;11(9):e0162765 [FREE Full text] [doi: 10.1371/journal.pone.0162765] [Medline: 27632172]
45. Menefee HK, Thompson MJ, Guterbock TM, Williams IC, Valdez RS. Mechanisms of communicating health information through facebook: implications for consumer health information technology design. J Med Internet Res 2016 Aug 11;18(8):e218 [FREE Full text] [doi: 10.2196/jmir.5949] [Medline: 27515151]

Abbreviations

CHD: coronary heart disease
CVD: cardiovascular disease
eHealth: electronic health
mHealth: mobile health
NSW: New South Wales
SD: standard deviation
SPSS: Statistical Package for the Social Sciences

Edited by G Eysenbach; submitted 22.05.17; peer-reviewed by J Rawstorn, A Conway; comments to author 06.07.17; revised version received 17.07.17; accepted 20.07.17; published 15.09.17.

Please cite as:
Partridge SR, Grunseit AC, Gallagher P, Freeman B, O'Hara BJ, Neubeck L, Due S, Paull G, Ding D, Bauman A, Phongsavan P, Roach K, Sadler L, Glinatsis H, Gallagher R
Cardiac Patients' Experiences and Perceptions of Social Media: Mixed-Methods Study
J Med Internet Res 2017;19(9):e323
URL: http://www.jmir.org/2017/9/e323/
doi:10.2196/jmir.8081
PMID:28916507

Original Paper

Blending Face-to-Face and Internet-Based Interventions for the Treatment of Mental Disorders in Adults: Systematic Review

Doris Erbe[1], Dipl Psych; Hans-Christoph Eichert[1], PhD; Heleen Riper[2], PhD; David Daniel Ebert[3], PhD

[1]Department of Special Education and Rehabilitation, University of Cologne, Cologne, Germany

[2]Department of Clinical Psychology, Vrije Universiteit Amsterdam, Amsterdam, Netherlands

[3]Department of Clinical Psychology and Psychotherapy, Friedrich-Alexander University Erlangen-Nuremberg, Erlangen, Germany

Corresponding Author:
David Daniel Ebert, PhD
Department of Clinical Psychology and Psychotherapy
Friedrich-Alexander University Erlangen-Nuremberg
Nägelsbachstr. 25a
Erlangen, 91052
Germany
Phone: 49 9131 85 67566
Fax: 49 9131 85 67576
Email: david.ebert@fau.de

Abstract

Background: Many studies have provided evidence for the effectiveness of Internet-based stand-alone interventions for mental disorders. A newer form of intervention combines the strengths of face-to-face (f2f) and Internet approaches (*blended interventions*).

Objective: The aim of this review was to provide an overview of (1) the different formats of blended treatments for adults, (2) the stage of treatment in which these are applied, (3) their objective in combining face-to-face and Internet-based approaches, and (4) their effectiveness.

Methods: Studies on blended concepts were identified through systematic searches in the MEDLINE, PsycINFO, Cochrane, and PubMed databases. Keywords included terms indicating face-to-face interventions ("inpatient," "outpatient," "face-to-face," or "residential treatment"), which were combined with terms indicating Internet treatment ("internet," "online," or "web") and terms indicating mental disorders ("mental health," "depression," "anxiety," or "substance abuse"). We focused on three of the most common mental disorders (depression, anxiety, and substance abuse).

Results: We identified 64 publications describing 44 studies, 27 of which were randomized controlled trials (RCTs). Results suggest that, compared with stand-alone face-to-face therapy, blended therapy may save clinician time, lead to lower dropout rates and greater abstinence rates of patients with substance abuse, or help maintain initially achieved changes within psychotherapy in the long-term effects of inpatient therapy. However, there is a lack of comparative outcome studies investigating the superiority of the outcomes of blended treatments in comparison with classic face-to-face or Internet-based treatments, as well as of studies identifying the optimal ratio of face-to-face and Internet sessions.

Conclusions: Several studies have shown that, for common mental health disorders, blended interventions are feasible and can be more effective compared with no treatment controls. However, more RCTs on effectiveness and cost-effectiveness of blended treatments, especially compared with nonblended treatments are necessary.

(J Med Internet Res 2017;19(9):e306) doi:10.2196/jmir.6588

KEYWORDS

mental health; Internet; psychotherapy; blended treatment

Introduction

Background

Empirical evidence suggests that Internet-based psychological interventions can be used to effectively treat adults, adolescents, and children for various mental disorders such as depression, anxiety, or problematic substance use [1-6]. Such interventions have several advantages over conventional face-to-face (f2f) interventions. For example, Internet-based interventions can be administered over long distances, may save therapists' time, allow both patients and clinicians to work at their own pace, save traveling time, and reduce the stigma of having a mental disorder or going to a psychologist or therapist [7,8].

On the other hand, Internet-based interventions may also have disadvantages when compared with face-to-face therapies. For instance, Internet interventions may require certain abilities such as computer and Internet skills, reading and writing skills, and, in comparison with traditional therapy settings, more self-reflection and eloquence when talking about one’s thoughts and feelings. Furthermore, it has been argued that this type of intervention may make it difficult for therapists to adequately react to crisis situations such as suicidality because nonverbal cues are missing as additional information when assessing whether dissociation of suicidal thoughts is possible [9]. Also, negative effects such as frustration due to failure or time pressure might be associated with the Internet treatment format [10-13]. These disadvantages often lead to stand-alone Internet treatments being regarded as low-threshold interventions for milder cases of mental disorders, whereas face-to-face therapy and pharmacotherapy are often regarded as options of choice for more severe symptoms [14]. In addition to guided or unguided Internet treatment as a stand-alone intervention, a newer treatment approach combines face-to-face sessions with Internet-based sessions into one integrated treatment. This approach is usually called *blended treatment* and aims at retaining the positive aspects associated with both forms of therapy while mitigating the disadvantages [15-17]. There are different potential advantages of blending Internet and face-to-face treatments. For example, viewed from a cost-effectiveness perspective, blended treatments could possibly diminish the number of face-to-face contacts and thereby decrease the overall costs of treatment. Blended treatments could also increase the effectiveness by, for example, increasing frequency from one to two sessions per week [18] through adding Internet sessions to face-to-face interventions, thereby increasing intensity and success without additional costs. Furthermore, adding Internet interventions might improve transfer to everyday life as Internet or mobile elements could be used to support behavior change during face-to-face sessions and thereby increase effectiveness of face-to-face psychotherapy. Blended format could potentially also reach individuals for whom either purely delivered Internet-based or pure face-to-face approaches are not a suitable treatment option and thereby increase the utilization of effective treatments.

As both face-to-face and Internet-based psychotherapy have advantages and disadvantages, combining the two approaches in a blended treatment might combine the best of two worlds.

Definition of Blended Interventions

As a clear definition of blended interventions is still missing [17], we define blended interventions in this study as treatment programs that use elements of both face-to-face and Internet-based interventions, including both the integrated and the sequential use of both treatment formats. Nonblended interventions comprise face-to-face treatments only or stand-alone Internet treatments. Within blended treatments, face-to-face contacts may be added to Internet interventions or, vice versa, the Internet-based part may be arranged as an adjunct to existing face-to-face programs. Internet interventions may also be arranged as an aftercare or maintenance element after acute phase treatment or as an early step within a stepped care program. *Stepped care* refers to a treatment program where interventions start with the least intensive and least costly treatment that is likely to work, progressing step-by-step with more intensive interventions for those patients insufficiently helped by the first or previous intervention. Blended stepped care treatments involve Internet-based treatments as one step within the sequence. See Textbox 1 for a subject index of types of blended interventions.

Textbox 1. Subject index

Nonblended interventions: Face-to-face (f2f) treatments or stand-alone Internet-treatments only

Blended interventions: Treatment programs that use elements of both face-to-face and Internet-based interventions, including sequential use of both forms of treatment

- Integrated blended interventions: Blended treatments where the Internet-based intervention part is arranged as an adjunct to face-to-face programs or vice versa so that face-to-face and Internet-based elements are provided within the same period. In integrated blended interventions, the focus can be either on the face-to-face treatment or on the Internet-based intervention.
- Sequential blended interventions: Blended treatments where the Internet-based intervention part is arranged before or after the face-to-face treatment such as within stepped care approaches or aftercare interventions that directly follow the face-to-face intervention.

Objective

This systematic review supplies an overview of research into blended interventions for mental health. Specifically, we focus on the following questions: (1) Which blended intervention concepts have been proposed in the researched literature in the treatment of common mental disorders (anxiety disorders, depression, and substance abuse)? (2) In which stage of treatment (such as first step, acute phase treatment, and maintenance phase) do the Internet interventions take place? (3) Which types of problem and target group do the blended interventions focus on? (4) What is the objective in combining face-to-face and Internet-based approaches? and (5) What evidence is there for the effectiveness of blended interventions?

Methods

Search Strategy

Studies of potential relevance were identified using a systematic search in the MEDLINE, PsycINFO, Cochrane, and PubMed databases. All studies up to December 2015 were included. Searches were performed using keywords indicating face-to-face interventions (“inpatient,” “outpatient,” “face to face,” or "residential treatment"), which were combined with terms indicating Internet treatment ("internet," "online," or "web") and terms indicating mental disorders ("mental health," "depression," "anxiety," or "substance abuse"). We focused on three of the most common mental disorders: depression, anxiety, and substance abuse. The bibliographies of the identified studies revealed additional sources.

Study Selection

Studies were included if they met the following inclusion criteria: (1) the study was on an intervention that was based on both an Internet and a face-to-face treatment element that was either integrated or delivered sequentially, (2) the study involved treatment for adults with depression, anxiety, or substance abuse, and (3) the study was published in English or German. Studies with mere self-help interventions were excluded.

After the initial database search and removal of duplicates, the title and abstract of the remaining studies were rated for the inclusion criteria independently by the first and second author. Interrater reliability was good (kappa, κ=.825, *P*<.001). The first and second authors then used the inclusion criteria to independently review all full-text publications judged as relevant in the title and abstract screening. Consensus through discussion was sought in cases of disagreement. If no agreement could be reached through discussion, the last author made a decisive judgment. This was the case for 2 studies.

In November 2016, we started a second search in the database of PubMed, searching for studies citing the studies we had found in the initial search. The extraction strategies as of the initial search were used for those studies, including the independent rating of studies by the first and second authors in the first and second steps.

Quality Assessment

New quality assessment criteria were created since there are no current guidelines for assessing the quality of blended intervention studies. The quality of each study was rated on five aspects: study design, randomization of study conditions, report of statistics, sample size (studies powered to detect effect sizes of a minimal important difference between blended and nonblended interventions of a priori defined Cohen *d* of 0.35 were classified as high quality), and existence of a nonblended active control group. Each aspect was rated on a 3-point scale (0-2), providing a score with a range from 0 to10 (see Table 1 for detailed descriptions of quality assessment). Studies obtaining a total score above two-thirds of the maximum score (ie, >7) were considered high quality studies.

Table 1. Study assessment criteria for scoring.

Aspect	Scoring
Study design	Which design did the study have? 2=controlled trial, 1=pre-post, 0=case study or unclear.
Randomization	Were participants randomized to conditions (depending on design)? 2=yes, 0=no.
Report of statistics	Are relevant statistics reported? 2=mean and standard deviation for outcome measures, effect sizes, and *P* values for significant differences are reported, 1=is lacking any of these, 0=lack several of these.
Sample size	Was the sample size adequate to detect effect sizes of *d*=0.35? 2=264 or more, 1=more than 132, 0=132 or less
Nonblended control group	Did the design involve a nonblended active control group? 2=nonblended control group with same number of sessions, 1=active control group without assured same number of sessions, 0=no active control group

Results

Study selection

See Figure 1 for a flowchart of included and excluded studies. In total, we identified 64 publications describing 44 studies. Eight of the included studies were study protocols that had not yet published results in November 2016.

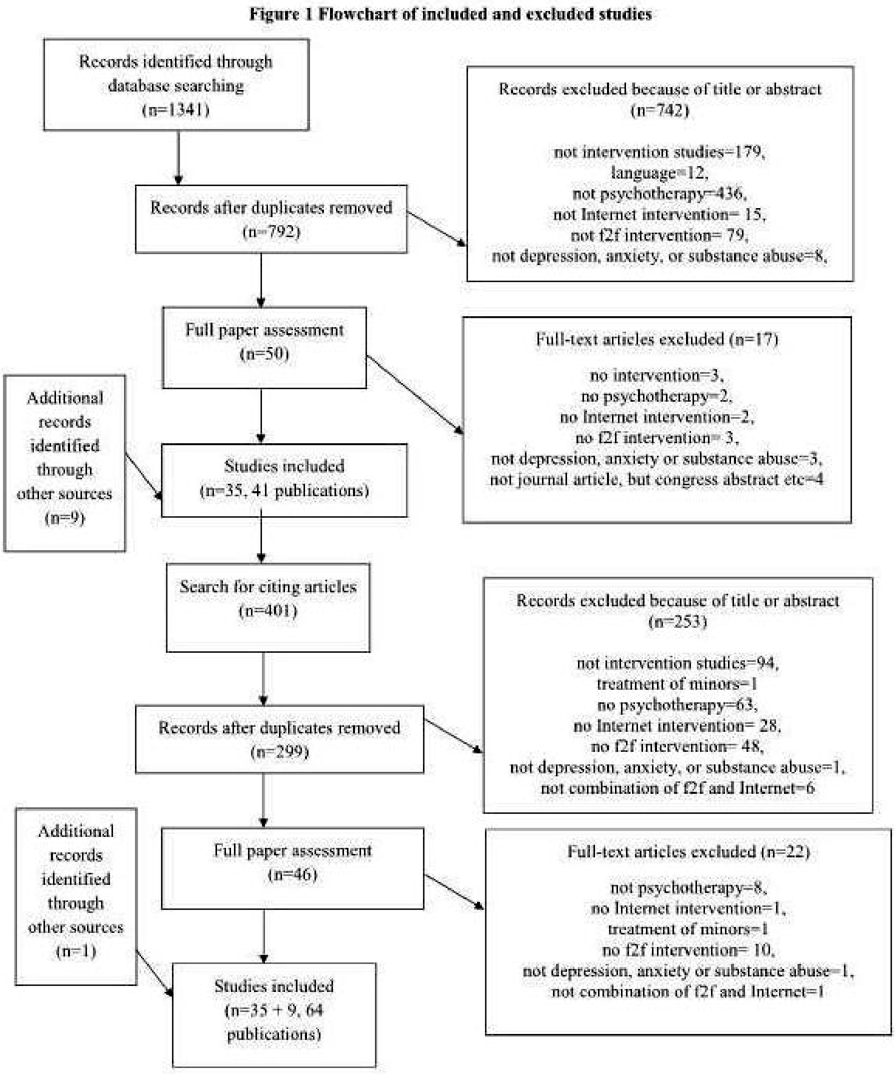
Of the included studies, 27 were randomized controlled trials (RCTs), 4 were non-RCTs, 5 were pre-post studies without a control group, 4 were case studies, 3 were preliminary evaluation or acceptability studies, and 1 was a qualitative study. See Multimedia Appendix 1 for detailed study characteristics.

Of the 44 studies, 8 were rated high quality studies (18%). A control group was involved in 31 (70%, 18/44) of the studies. Twelve studies (27%, 12/44) involved a nonblended active control group with the same number of sessions as the intervention group. Only 6 studies (14%, 6/44) involved a sample of 264 or more. All relevant statistics, including effect sizes, were reported by 17 studies (39%, 17/44). See Multimedia Appendix 3 for all quality assessment scores.

Most studies were conducted or planned in the United States (n=12), followed by the Netherlands (n=9) and Germany (n=8). Six studies were conducted in Australia, four in Norway, three in the United Kingdom, and two in Sweden.

The Internet part of the studies' blended interventions generally used Web-based programs with modules combining techniques such as cognitive, behavioral, and/or emotion-focused interventions, some of them with email support. However, there were two exceptions. One study [23] reports on using emails and Internet sites on psychoeducation and other issues to improve organization and to assist with common problems related to depression. The other study [44] reports offering weekly Internet group chats with a therapist (in open groups of 8-10 participants) to focus on problems arising in readjusting to everyday life after inpatient treatment.

Figure 1. Flowchart of included and excluded studies.



Disorders Addressed

Of the included 44 studies, 20 studies focused on treating depression only, eight focused on anxiety disorders only, and eight focused on substance abuse only. One study described the treatment of comorbid depression and substance abuse; three studies treated both depression and anxiety. The remaining four studies had a transdiagnostic concept involving depression, anxiety, and other mental disorders. See Multimedia Appendix 1 for details on disorders addressed.

Concept of Blended Care and Stage of Treatment

Following the study selection process, we clustered the studies into the following types of blended care (Multimedia Appendix 1):

1. Integrated blended interventions with face-to-face focus: These blended interventions are based on an face-to-face intervention that is complemented or partly replaced by an Internet intervention; face-to-face and Internet-based elements are provided within the same period.
2. Integrated blended interventions with Internet focus: These blended interventions are based on Internet interventions that are partly replaced or complemented by face-to-face sessions; face-to-face and Internet-based elements are provided within the same period.
3. Sequential blended interventions with Internet, then face-to-face: These blended interventions arrange the Internet intervention part before the face-to-face treatment, such as within stepped care.
4. Sequential blended interventions with face-to-face, then Internet: These blended interventions arrange the Internet intervention part after the face-to-face treatment as in an aftercare program.

The majority of studies (n=29) used a concept of integrated blended intervention, with face-to-face and Internet-based elements being provided within the same period. Among those, 18 studies focused on the face-to-face intervention, considering the Internet intervention as a replacement of some of the face-to-face sessions or as an adjunct, whereas 11 studies focused on the Internet intervention as the basis of treatment where the face-to-face sessions served as an adjunct, for example, for increasing adherence to the Internet-based modules.

The remaining 15 studies presented sequential blended interventions, arranging the Internet intervention part before or following the face-to-face treatment. Nine of them placed the Internet intervention before the face-to-face treatment either as part of a stepped or matched care program (n=4) or for bridging waiting time for referrals on waiting lists for face-to-face psychotherapy (n=5). Six studies placed the Internet intervention after the face-to-face treatment as an aftercare concept.

Aims

The studies' aims for choosing to use a blended intervention can be classified by the concepts that were used (see Multimedia Appendix 1 for details). As some studies mentioned various aims, we clustered the interventions by their main objective. See Figure 2 for a summary of aims of blended therapy.

Integrated Blended Interventions With Face-to-Face (F2F) Focus

Among the 18 integrated blended interventions with face-to-face focus, seven aimed at delegating some elements of face-to-face therapy to Internet-based cognitive behavioral therapy (iCBT) and thereby, saving clinician time and reducing overall costs [15,51,54,58,62,69,75]. Nine of them aimed at supporting face-to-face therapy by delivering additional Internet elements and thereby increasing effectiveness of face-to-face psychotherapy [23-25,37,40,55,60,61,77,78]. Two studies stated as their aim integrating Internet elements with face-to-face psychotherapy to establish a proactive and long-term approach to the management of chronic conditions, thus providing long-term support for patients with chronic or recurrent mental diseases beyond the acute phase of face-to-face treatment [59,68].

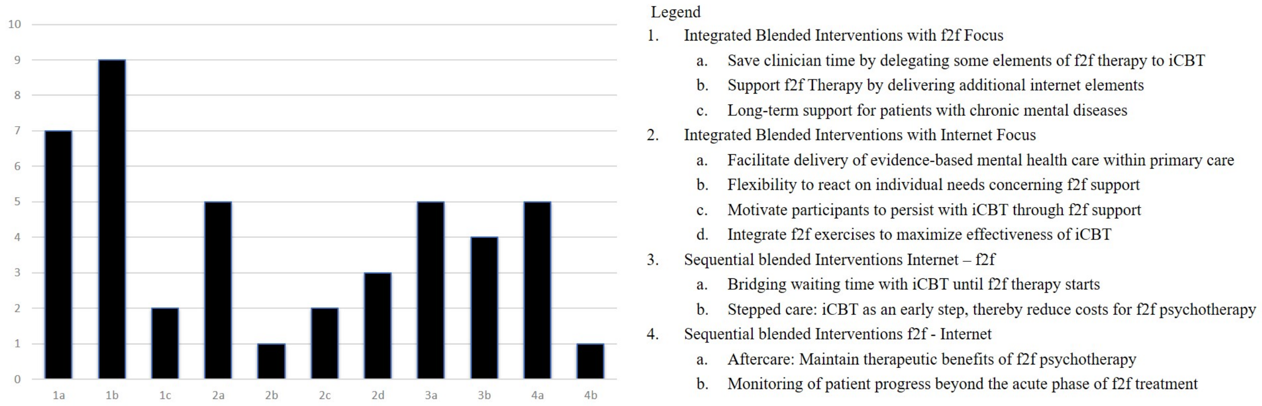
Integrated Blended Interventions With Internet Focus

Five of the 11 integrated blended interventions with an Internet focus aimed at improving the delivery of evidence-based treatment in primary care [31,48,50,65,72], for instance, by assisting GPs in providing evidence-based mental health care programs [65]. Three of the studies aimed at integrating face-to-face sessions to maximize the effectiveness of iCBT [19,64,70]. Other aims of these studies were to offer an Internet intervention to all participants with complementary face-to-face sessions as needed by the individual participant and as such to increase flexibility to meet the needs of different clients concerning face-to-face support [79], or, through face-to-face support, to motivate participants to persist with iCBT [45,74].

Sequential Blended Interventions With Internet, Then F2F

The nine sequential blended interventions that started with the Internet intervention either aimed at bridging waiting time (n=5) with iCBT until face-to-face therapy started [53,57,67,73,80] or worked with a blended stepped care concept, aiming at delivering low-threshold iCBT as an early step and thereby reducing costs for subsequent face-to-facepsychotherapy (n=4) (by reducing either the number of v sessions or the number of patients treated in the subsequent face-to-facephase) [20,26,46,47].

Figure 2. Aims of blended therapy.



Sequential Blended Interventions With F2F, Then Internet

Five of the six sequential blended interventions that started with the face-to-face intervention were designed as aftercare programs aiming at maintaining therapeutic benefits of face-to-faceresidential or inpatient psychotherapy through subsequent iCBT [22,41,44,56,76]. Another study aimed to monitor patient progress beyond the acute phase of face-to-face treatment for a long time and to maintain the therapeutic relationship in the absence of face-to-face contacts in a remote setting [71].

Outcome

See Multimedia Appendix 2 for details. Given the variety of different study types (eg, study protocols, case studies, and qualitative studies) and that outcome measures, study designs, and aims differ substantially, it is not currently possible to summarize the effects using meta-analytic techniques.

Of the 44 identified studies, eight were study protocols and 36 had been completed. Among the 36 completed studies, 23 involved a control group. Out of the 23 completed studies with a control group, four studies compared the blended intervention with no treatment, for instance, a waiting-list group [19,41,44,48]. Nine studies compared the blended intervention group with a nonblended face-to-face intervention with an equal amount or more sessions [24,32,50,53-55,57,62,75], six studies compared the blended intervention with treatment-as-usual without controlling the number of sessions [20,26,32,37,40], and two studies compared the blended intervention with a nonblended Internet intervention [64,65]. The remaining two studies involved both Internet-based and face-to-face nonblended control groups [55,70].

Among the eight study protocols, two studies compared the blended intervention with TAU without controlling for the number of sessions [46,51,59], whereas four studies compared the blended intervention-group with a nonblended face-to-face intervention with an equal number or more sessions [15,53,60,69]. One study compared the blended intervention (in that case, aftercare after inpatient treatment) with a placebo (inpatient treatment plus access after discharge to an Internet platform with information on depression [76]), and one study compared it with no treatment [67].

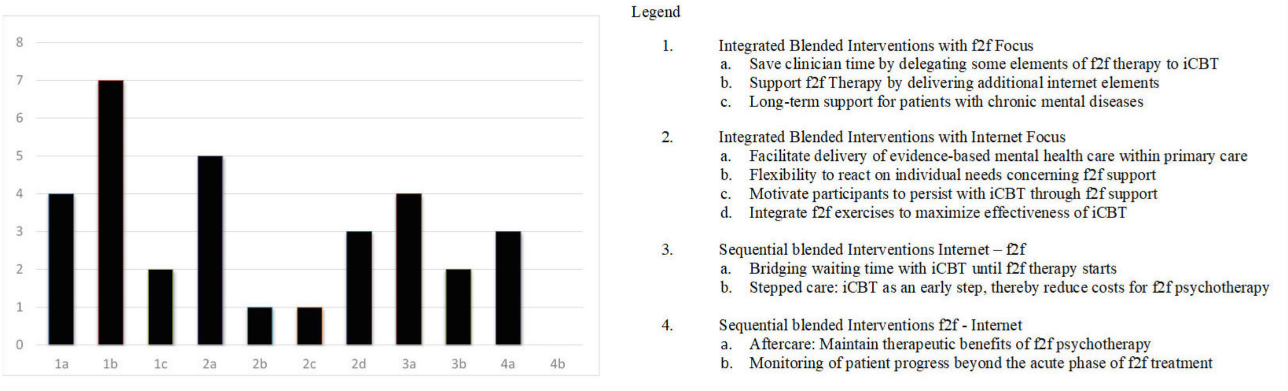
Although cost-effectiveness or cost-savings was in some way considered by almost all 44 studies, only three of the completed studies elaborated on it and evaluated potential cost-effectiveness or cost-savings [33,52,81]. However, three additional study protocols claimed to be planning to evaluate cost-effectiveness [15,46,51,69]. See Figure 3 for a summary of aims of blended therapy in completed studies.

Integrated Blended Interventions With F2F Focus

Of the six studies that were delegating some elements of face-to-face therapy to iCBT, three studies were able to show that, by doing that, 50% to 86 % of clinician time could be saved without reducing the therapeutic outcome of depression and anxiety treatment [54,62,75]. The other three were study protocols [15,51,69].

Six of the nine studies aiming at supporting face-to-face therapy by delivering Internet elements were able to show that adding Internet elements can lead to lower dropout rates and/or greater abstinence rates of patients with substance abuse compared with stand-alone face-to-face interventions [24,40,55,77]. One study had a pre-post design [61] and showed that its blended concept led to a reduction in symptomatology that was maintained for 12 months. One was a case study where a patient diagnosed with depression as well as her therapist were able to make creative use of Web-based resources for purposes such as psychoeducation or job search [23]. One was a study protocol [60].

Figure 3. Aims of blended therapy of studies with published outcome.



One study aiming at establishing a proactive and long-term approach to the management of chronic mental diseases beyond the acute phase of face-to-face treatment [68] showed large and significant pre-post effect sizes. One study with a similar aim [59] showed that an Internet-delivered adjunct to TAU may reduce the number of unwell weeks in patients with recurrent depression compared with TAU and thus, reduce the lifelong burden of depression.

Integrated Blended Interventions With Internet Focus

Regarding the aim of integrating face-to-face sessions into Internet interventions to maximize effectiveness of iCBT, Sethi et al [70] found the blended intervention to be superior in reducing symptoms of depression, anxiety, and/or automatic negative thoughts in comparison with three control groups: nonblended Internet intervention, nonblended face-to-faceintervention, and no intervention. One study treating depression [64] did not find a difference between the blended and the nonblended intervention, whereas a third study did not have a nonblended control group but found the blended intervention to be superior compared with a waiting-list control group [19].

Using blended interventions to improve the delivery of evidence-based treatment in primary care has been successful in acceptability and in reducing symptomatology [31,48,50,65,72]. However, two controlled studies [50,65] could not show a superiority of the blended intervention over an active nonblended control group.

In a study aiming at increasing the flexibility of an Internet intervention by offering complementary face-to-face sessions as needed by the individual participant [79], the average number of face-to-face sessions participants needed was equal to or less than 3.7, with the program still substantially and significantly reducing depression symptomatology. This was a pilot study without a control group, yet the authors cite a study that treated participants with similar pre- and post-Beck depression inventory scores as their study [82]. The other study's participants needed an average of 11.6 face-to-face sessions for a similar symptom reduction, and the authors therefore claim that their treatment needs about eight fewer individual sessions per client.

A study aiming at motivating participants to persist with iCBT through face-to-face support [74] used a qualitative design. It found that for persistence in a blended treatment, acknowledgment may be related to flexibility and feedback from a qualified therapist in the face-to-face consultations beneath personal resources such as a sense of belonging toward partners, family, and friends. A case study with a similar aim [45] showed good adherence and symptom remission in a woman suffering from antepartum depression, who was thus able to avoid antidepressant use during pregnancy.

Sequential Blended Interventions With Internet, Then F2F

Of the five studies aiming at bridging waiting time with iCBT until face-to-face therapy starts, two studies [53,57] showed significant between-group effect sizes between the iCBT waiting-list group and the control waiting-list group. One study failed to show a superiority of the iCBT waiting-list group compared with a control condition with participants who remained on the wait-list with a self-help booklet [83]. One of the studies showed in a pre-post design that the Internet intervention could substantially and significantly reduce symptomatology [73]. The remaining study with this aim [67] is a study protocol.

Two of the completed studies working with a blended stepped care concept did not find a significant superiority of the blended intervention compared with face-to-face [47] or TAU [20]. Haug et al [47] even found a significantly better outcome of the control group (nonblended f2f) in one of the outcome measures. Neither Haug et al nor Braamse et al [20] reported results on cost-saving or cost-effectiveness; however, one study protocol with a stepped care concept mentioned planning to evaluate cost-effectiveness [46].

Sequential Blended Interventions With F2F, Then Internet

Two of the five studies with Internet intervention designed as maintenance treatment found a substantially and significantly lower relapse rate compared with access to TAU groups [41,44]. One further study was a case study reporting how iCBT aftercare substantially reduced depression, leading to remission after discharge from inpatient treatment [22]. One study with an empirical correlational design and without a control group showed a significant relationship between the number of iCBT modules accessed and substance abuse reduction in the year following inpatient treatment when controlling for motivation, self-efficacy, and pretreatment substance abuse [56]. One study was a study protocol [76].

One case study aiming at long-term monitoring of patient progress beyond the acute phase of face-to-face treatment, as well as maintaining the therapeutic relationship in the absence of face-to-face contacts in a remote setting [71], reported continuing therapeutic contacts and a reduction from very severe depression symptoms to remission within 46 weeks.

Discussion

Principal Findings

This study has shown that, in the past few years, a growing number of blended interventions that combine Internet and face-to-face interventions have been developed for common mental health problems. The interventions we found have different concepts and various aims. First results are encouraging and suggest that, compared with stand-alone face-to-face interventions, blended therapy may save clinician time without reducing therapy outcome, can lead to lower dropout rates and/or greater abstinence rates of patients with substance abuse, may help maintain effects of inpatient therapy, and may even increase the effects of psychotherapy, although results are mixed and more research is clearly needed.

Compared with the field of both Internet-based stand-alone treatment and face-to-face interventions, the field of blended interventions is, however, under development and still small. Most aims of the interventions stated in the studies have not been evaluated rigorously. For instance, only 19 out of 36 completed studies were RCTs, and only eight of them ensured comparability by involving a nonblended intervention control group with the same number of sessions [27,50,57,62,70,75,78,83]. Only eight studies were considered high quality studies, and only six studies involved a control group big enough to detect effect sizes of d=0.35 or less. Also, the cost-effectiveness of blended treatments compared with face-to-face psychotherapy has only been evaluated in three out of the 36 completed studies. However, this issue is focused upon in current large research projects [46,84].

Several questions remain. For instance, not much is known about the optimal ratio of Internet and face-to-facesessions that would allow costs to be minimized while maintaining or increasing effectiveness. Only one study in our review is moving toward answering this question: In the study by Jacmon et al [79], clients received face-to-face sessions as wished as an adjunct to iCBT. Also, clients were actively invited by therapists to undertake face-to-face contacts if a depression measure indicated that depression continued at clinical levels. Jacmon et al reported that this intervention produced effects at least as large as a completely face-to-face treatment, with an average of about eight fewer individual sessions per client in comparison with a similar nonblended face-to-face treatment in another study. However, the findings of this pilot study cannot be generalized because of its small sample size and lack of a control group. Another open question is which elements of face-to-face therapy can most suitably be delegated to the Internet. Possibly, elements that do not need an intensive dialogue between client and therapist, such as psychoeducation, may most easily be delegated to the Internet [16]. Also, it remains unclear who would benefit most from which relationship between face-to-face and Internet modules. For instance, patients who have difficulty in expressing all of their thoughts and feelings in writing (eg, because of lack of introspection or because they are not conscious of their own role in maintaining their disorder) might need more elements of face-to-face intervention psychotherapy. This might also be the case for patients with severe, chronic, or personality disorders. A helpful tool to set up a personalized blended treatment taking the patients' characteristics into account might be the recently developed "Fit for Blended Care" instrument [17].

Moreover, although a number of studies stated that increasing the effectiveness of face-to-face psychotherapy is an aim of the blending of treatments with Internet options, evidence for this hypothesis is limited to four studies with an integrated concept [24,25,32,70]. Such an assumption is, however, supported by a recent meta-analysis by Lindhiem and colleagues [85]. On the basis of 10 RCTs, they found that, in comparison with strictly on-site interventions, psychological interventions were considerably more effective for a range of conditions when behavior changes between face-to-face sessions had been supported by a mobile component, such as short message service (SMS; *d*=0.27). There is also some evidence that telephone-based interventions may help maintain advances achieved during outpatient treatment or increase effectiveness of face-to-face interventions [86,87]. Nevertheless, more research is clearly needed to conclude whether the effectiveness of psychotherapy can potentially be increased by using blended treatments. Besides the identified advantages of blended interventions, for example, saving clinician time, improving success rates in the reduction of symptomatology, and helping to prevent relapse after face-to-face therapy, these interventions might possibly have limitations compared with face-to-face interventions. For instance, in the study by Marks et al [62], dropout was substantially higher in the mainly Internet-based group compared with the face-to-face control group (43% if mainly computer-guided and 24% if entirely clinician-guided). On the other hand, other studies have reported similar [50,75] or even lower [24] dropout rates of a blended compared with a nonblended intervention. Yet, the questions of blended interventions’ negative effects and potential losses compared with nonblended interventions remain to be answered, just as much as the one about potential negative effects of Internet-based treatments [12].

An interesting question we encountered is what blended treatments exactly entail. The identified studies in our review that explicitly use the term “blended” are the more recently published ones and describe integrated blended treatments, that is, combined treatments that provide face-to-face and Internet-based elements within the same period. Although there is currently no clear definition, this description might be an implicit understanding of blended interventions that is more limiting than our definition. We consider that blended interventions should also include the combination of the Internet-based interventions arranged before or after the face-to-face treatments following clear rules and procedures, such as within stepped care or aftercare that directly follows the acute phase treatment. In light of the numerous studies describing sequential blended interventions, we decided to use this wider definition.

Limitations

This study has some limitations. As within every systematic review, the risk of selection bias when including relevant studies needs to be considered. However, our use of independent ratings by two of the authors worked against this bias. In addition, publication bias needs to be considered. We did not contact authors for additional data or additional studies, which would have automatically limited the number of studies that could be included. Also, we only reviewed bibliographies from included studies, so we possibly missed studies that were cited in papers that we reviewed but did not include. Furthermore, the types of studies we included were heterogeneous (for instance, we included study protocols). It is possible that more narrow inclusion criteria (eg, randomized controlled studies with nonblended active control groups only) would have produced more information about the effectiveness of blended interventions. However, the number of such studies in the field is yet very small, and future studies are needed to explore whether blended treatments can, for example, be superior compared with nonblended treatments with regard to effect sizes or lower costs.

For further research, it would be of interest to explore effectiveness and cost-effectiveness of blended concepts, especially concerning the optimal balance of face-to-face and Internet interventions. Information on this aspect would help determine which therapy with which theoretical foundation (such as cognitive behavioral therapy or psychoanalysis) is feasible for blended interventions, as well as where the optimal balance of therapy modules lies for individual patients in light of the type and severity of disorder, state of motivation, ability of introspection, and other variables such as age, gender, and computer skills.

Conclusions

To conclude, we have found that several studies have shown that blended interventions are feasible and effective compared with no treatment controls. There are many different kinds of blended concepts that, in every phase of treatment, may offer added value concerning either effectiveness or cost-effectiveness. However, to evaluate the actual benefit of blended concepts for mental health care, more RCTs on effectiveness and cost-effectiveness compared with traditional nonblended psychotherapy are required. Thus, more research is needed, especially concerning disorders for which blended interventions are particularly effective, the amount of face-to-face contact that is needed, and the parts of therapy that can be delegated to the Internet.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Previous psychological research combining Internet and face-to-face psychotherapy.

[PDF File (Adobe PDF File), 42KB - jmir_v19i9e306_app1.pdf]

Multimedia Appendix 2

Outcome of previous psychological research combining Internet and face-to-face (f2f) psychotherapy.

[PDF File (Adobe PDF File), 87KB - jmir_v19i9e306_app2.pdf]

Multimedia Appendix 3

Study quality assessment of blended treatments.

[PDF File (Adobe PDF File), 42KB - jmir_v19i9e306_app3.pdf]

References

1. Andersson G, Topooco N, Havik O, Nordgreen T. Internet-supported versus face-to-face cognitive behavior therapy for depression. Expert Rev Neurother 2016;16(1):55-60. [doi: 10.1586/14737175.2015.1125783] [Medline: 26610160]
2. Ebert DD, Zarski AC, Christensen H, Stikkelbroek Y, Cuijpers P, Berking M, et al. Internet and computer-based cognitive behavioral therapy for anxiety and depression in youth: a meta-analysis of randomized controlled outcome trials. PLoS One 2015;10(3):e0119895 [FREE Full text] [doi: 10.1371/journal.pone.0119895] [Medline: 25786025]
3. Mayo-Wilson E, Montgomery P. Media-delivered cognitive behavioural therapy and behavioural therapy (self-help) for anxiety disorders in adults. Cochrane Database Syst Rev 2013(9):1-48. [doi: 10.1002/14651858.CD005330.pub4] [Medline: 24018460]
4. 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]
5. Riper H, Blankers M, Hadiwijaya H, Cunningham J, Clarke S, Wiers R, et al. Effectiveness of guided and unguided low-intensity internet interventions for adult alcohol misuse: a meta-analysis. PLoS One 2014;9(6):e99912 [FREE Full text] [doi: 10.1371/journal.pone.0099912] [Medline: 24937483]
6. Königbauer J, Letsch J, Ebler P, Ebert DD, Baumeister H. Internet- and mobile-based depression interventions for people with diagnosed depression: a systematic review and meta-analysis. J Affect Disord 2017 Jul 10;223:28-40. [doi: 10.1016/j.jad.2017.07.021] [Medline: 28715726]
7. Cuijpers P, van SA, Andersson G. Internet-administered cognitive behavior therapy for health problems: a systematic review. J Behav Med 2008 Apr;31(2):169-177 [FREE Full text] [doi: 10.1007/s10865-007-9144-1] [Medline: 18165893]
8. Ebert DD, Berking M, Thiart H, Riper H, Laferton JA, Cuijpers P, et al. Restoring depleted resources: Efficacy and mechanisms of change of an internet-based unguided recovery training for better sleep and psychological detachment from work. Health Psychol 2015 Dec;34S:1240-1251. [doi: 10.1037/hea0000277] [Medline: 26651465]
9. Berger T, Andersson G. [Internet-based psychotherapies: characteristics and empirical evidence]. Psychother Psychosom Med Psychol 2009;59(3-4):159-66; quiz 167. [doi: 10.1055/s-0028-1090162] [Medline: 19347788]
10. Boettcher J, Rozental A, Andersson G, Carlbring P. Side effects in internet-based interventions for social anxiety disorder. Internet Interv 2014 Mar;1(1):3-11. [doi: 10.1016/j.invent.2014.02.002]
11. Rozental A, Boettcher J, Andersson G, Schmidt B, Carlbring P. Negative effects of internet interventions: a qualitative content analysis of patients' experiences with treatments delivered online. Cogn Behav Ther 2015;44(3):223-236. [doi: 10.1080/16506073.2015.1008033] [Medline: 25705924]
12. Rozental A, Andersson G, Boettcher J, Ebert DD, Cuijpers P, Knaevelsrud C, et al. Consensus statement on defining and measuring negative effects of Internet interventions. Internet Interv 2014 Mar;1(1):12-19. [doi: 10.1016/j.invent.2014.02.001]
13. 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. [doi: 10.1017/S0033291716001562] [Medline: 27649340]
14. Moritz S, Schröder J, Meyer B, Hauschildt M. The more it is needed, the less it is wanted: attitudes toward face-to-face intervention among depressed patients undergoing online treatment. Depress Anxiety 2013 Feb;30(2):157-167. [doi: 10.1002/da.21988] [Medline: 22930656]
15. 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 Oct 18;14:290 [FREE Full text] [doi: 10.1186/s12888-014-0290-z] [Medline: 25326035]
16. van der Vaart R, Witting M, Riper H, Kooistra L, Bohlmeijer ST, van Gemert-Pijnen LJ. Blending online therapy into regular face-to-face therapy for depression: content, ratio and preconditions according to patients and therapists using a Delphi study. BMC Psychiatry 2014 Dec 14;14:355 [FREE Full text] [doi: 10.1186/s12888-014-0355-z] [Medline: 25496393]
17. Wentzel J, van der Vaart R, Bohlmeijer ET, van der Vaart R. Mixing online and face-to-face therapy: how to benefit from blended care in mental health care. JMIR Ment Health 2016 Feb 09;3(1):e9 [FREE Full text] [doi: 10.2196/mental.4534] [Medline: 26860537]
18. 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 Jul;149(1-3):1-13. [doi: 10.1016/j.jad.2013.02.030] [Medline: 23528438]

19. Andersson G, Carlbring P, Holmström A, Sparthan E, Furmark T, Nilsson-Ihrfelt E, et al. Internet-based self-help with therapist feedback and in vivo group exposure for social phobia: a randomized controlled trial. J Consult Clin Psychol 2006 Aug;74(4):677-686. [doi: 10.1037/0022-006X.74.4.677] [Medline: 16881775]
20. Braamse AM, van MB, Visser OJ, Boenink AD, Cuijpers P, Eeltink CE, et al. A randomized clinical trial on the effectiveness of an intervention to treat psychological distress and improve quality of life after autologous stem cell transplantation. Ann Hematol 2016 Jan;95(1):105-114 [FREE Full text] [doi: 10.1007/s00277-015-2509-6] [Medline: 26420062]
21. Braamse AM, van Meijel B, Visser OJ, Boenink AD, Cuijpers P, Eeltink CE, et al. A randomized clinical trial on the effectiveness of an intervention to treat psychological distress and improve quality of life after autologous stem cell transplantation. Ann Hematol 2016 Jan;95(1):105-114 [FREE Full text] [doi: 10.1007/s00277-015-2509-6] [Medline: 26420062]
22. Brakemeier E, Marchner J, Gutgsell S, Engel V, Radtke M, Tuschen-Caffier B, et al. CBASP@home: ein internetbasiertes situationsanalysen- training zur stabilisierung des therapieerfolgs nach stationärer therapie für chronisch depressive patienten. Verhaltenstherapie 2013 Aug;23:190-203. [doi: 10.1159/00035481]
23. Callan JA, Howland RH, Puskar K. Using computers and the Internet for psychiatric nursing intervention. J Psychosoc Nurs Ment Health Serv 2009 Jan;47(1):13-14. [Medline: 19227104]
24. Campbell AN, Nunes EV, Matthews AG, Stitzer M, Miele GM, Polsky D, et al. Internet-delivered treatment for substance abuse: a multisite randomized controlled trial. Am J Psychiatry 2014 Jun;171(6):683-690 [FREE Full text] [doi: 10.1176/appi.ajp.2014.13081055] [Medline: 24700332]
25. Cochran G, Stitzer M, Campbell AN, Hu MC, Vandrey R, Nunes EV. Web-based treatment for substance use disorders: differential effects by primary substance. Addict Behav 2015 Jun;45:191-194 [FREE Full text] [doi: 10.1016/j.addbeh.2015.02.002] [Medline: 25697725]
26. 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 07;192(11 Suppl):S36-S39. [Medline: 20528706]
27. Campbell AN, Nunes EV, Pavlicova M, Hatch-Maillette M, Hu MC, Bailey GL, et al. Gender-based outcomes and acceptability of a computer-assisted psychosocial intervention for substance use disorders. J Subst Abuse Treat 2015 Jun;53:9-15 [FREE Full text] [doi: 10.1016/j.jsat.2014.12.006] [Medline: 25613105]
28. Cunningham C, Stitzer M, Campbell AN, Pavlicova M, Hu MC, Nunes EV. Contingency management abstinence incentives: cost and implications for treatment tailoring. J Subst Abuse Treat 2017 Jan;72:134-139. [doi: 10.1016/j.jsat.2015.08.010] [Medline: 26482136]
29. Tofighi B, Campbell AN, Pavlicova M, Hu MC, Lee JD, Nunes EV. Recent internet use and associations with clinical outcomes among patients entering addiction treatment involved in a web-delivered psychosocial intervention study. J Urban Health 2016 Oct;93(5):871-883. [doi: 10.1007/s11524-016-0077-2] [Medline: 27653383]
30. Campbell AN, Nunes EV, Miele GM, Matthews A, Polsky D, Ghitza UE, et al. Design and methodological considerations of an effectiveness trial of a computer-assisted intervention: an example from the NIDA Clinical Trials Network. Contemp Clin Trials 2012 Mar;33(2):386-395 [FREE Full text] [doi: 10.1016/j.cct.2011.11.001] [Medline: 22085803]
31. Campbell AN, Turrigiano E, Moore M, Miele GM, Rieckmann T, Hu MC, et al. Acceptability of a web-based community reinforcement approach for substance use disorders with treatment-seeking American Indians/Alaska Natives. Community Ment Health J 2015 May;51(4):393-403 [FREE Full text] [doi: 10.1007/s10597-014-9764-1] [Medline: 25022913]
32. Carroll KM, Ball SA, Martino S, Nich C, Babuscio TA, Nuro KF, et al. Computer-assisted delivery of cognitive-behavioral therapy for addiction: a randomized trial of CBT4CBT. Am J Psychiatry 2008 Jul;165(7):881-888 [FREE Full text] [doi: 10.1176/appi.ajp.2008.07111835] [Medline: 18450927]
33. Olmstead TA, Ostrow CD, Carroll KM. Cost-effectiveness of computer-assisted training in cognitive-behavioral therapy as an adjunct to standard care for addiction. Drug Alcohol Depend 2010 Aug 01;110(3):200-207 [FREE Full text] [doi: 10.1016/j.drugalcdep.2010.02.022] [Medline: 20392575]
34. Carroll KM, Ball SA, Martino S, Nich C, Babuscio TA, Rounsaville BJ. Enduring effects of a computer-assisted training program for cognitive behavioral therapy: a 6-month follow-up of CBT4CBT. Drug Alcohol Depend 2009 Feb 01;100(1-2):178-181 [FREE Full text] [doi: 10.1016/j.drugalcdep.2008.09.015] [Medline: 19041197]
35. Carroll KM, Kiluk BD, Nich C, Babuscio TA, Brewer JA, Potenza MN, et al. Cognitive function and treatment response in a randomized clinical trial of computer-based training in cognitive-behavioral therapy. Subst Use Misuse 2011;46(1):23-34 [FREE Full text] [doi: 10.3109/10826084.2011.521069] [Medline: 21190403]
36. Sugarman DE, Nich C, Carroll KM. Coping strategy use following computerized cognitive-behavioral therapy for substance use disorders. Psychol Addict Behav 2010 Dec;24(4):689-695 [FREE Full text] [doi: 10.1037/a0021584] [Medline: 21198228]
37. Carroll KM, Kiluk BD, Nich C, Gordon MA, Portnoy GA, Marino DR, et al. Computer-assisted delivery of cognitive-behavioral therapy: efficacy and durability of CBT4CBT among cocaine-dependent individuals maintained on methadone. Am J Psychiatry 2014 Apr;171(4):436-444 [FREE Full text] [doi: 10.1176/appi.ajp.2013.13070987] [Medline: 24577287]
38. Carroll KM, Herman A, DeVito EE, Frankforter TL, Potenza MN, Sofuoglu M. Catehol-o-methyltransferase gene Val158met polymorphism as a potential predictor of response to computer-assisted delivery of cognitive-behavioral therapy among

cocaine-dependent individuals: preliminary findings from a randomized controlled trial. Am J Addict 2015 Aug;24(5):443-451 [FREE Full text] [doi: 10.1111/ajad.12238] [Medline: 25930952]
39. Morie KP, Nich C, Hunkele K, Potenza MN, Carroll KM. Alexithymia level and response to computer-based training in cognitive behavioral therapy among cocaine-dependent methadone maintained individuals. Drug Alcohol Depend 2015 Jul 01;152:157-163 [FREE Full text] [doi: 10.1016/j.drugalcdep.2015.04.004] [Medline: 25982006]
40. Christensen DR, Landes RD, Jackson L, Marsch LA, Mancino MJ, Chopra MP, et al. Adding an Internet-delivered treatment to an efficacious treatment package for opioid dependence. J Consult Clin Psychol 2014 Dec;82(6):964-972 [FREE Full text] [doi: 10.1037/a0037496] [Medline: 25090043]
41. Ebert D, Tarnowski T, Gollwitzer M, Sieland B, Berking M. A transdiagnostic internet-based maintenance treatment enhances the stability of outcome after inpatient cognitive behavioral therapy: a randomized controlled trial. Psychother Psychosom 2013;82(4):246-256. [doi: 10.1159/000345967] [Medline: 23736751]
42. Ebert DD, Gollwitzer M, Riper H, Cuijpers P, Baumeister H, Berking M. For whom does it work? moderators of outcome on the effect of a transdiagnostic internet-based maintenance treatment after inpatient psychotherapy: randomized controlled trial. J Med Internet Res 2013 Oct 10;15(10):e191 [FREE Full text] [doi: 10.2196/jmir.2511] [Medline: 24113764]
43. Ebert DD, Hannig W, Tarnowski T, Sieland B, Götzky B, Berking M. [Web-based rehabilitation aftercare following inpatient psychosomatic treatment]. Rehabilitation (Stuttg) 2013 Jun;52(3):164-172. [doi: 10.1055/s-0033-1345191] [Medline: 23761205]
44. Golkaramnay V, Bauer S, Haug S, Wolf M, Kordy H. The exploration of the effectiveness of group therapy through an Internet chat as aftercare: a controlled naturalistic study. Psychother Psychosom 2007;76(4):219-225. [doi: 10.1159/000101500] [Medline: 17570960]
45. Hantsoo L, Epperson CN, Thase ME, Kim DR. Antepartum depression: treatment with computer-assisted cognitive-behavioral therapy. Am J Psychiatry 2013 Aug;170(8):929-930 [FREE Full text] [doi: 10.1176/appi.ajp.2013.13020254] [Medline: 23903340]
46. Härter M, Heddaeus D, Steinmann M, Schreiber R, Brettschneider C, König HH, et al. [Collaborative and stepped care for depression: Development of a model project within the hamburg network for mental health (psychenet.de)]. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 2015 Apr;58(4-5):420-429. [doi: 10.1007/s00103-015-2124-7] [Medline: 25698121]
47. Haug T, Nordgreen T, Öst LG, Kvale G, Tangen T, Andersson G, et al. Stepped care versus face-to-face cognitive behavior therapy for panic disorder and social anxiety disorder: predictors and moderators of outcome. Behav Res Ther 2015 Aug;71:76-89. [doi: 10.1016/j.brat.2015.06.002] [Medline: 26081010]
48. 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 Aug 05;15(8):e153 [FREE Full text] [doi: 10.2196/jmir.2714] [Medline: 23916965]
49. Høifødt RS, Mittner M, Lillevoll K, Katla SK, Kolstrup N, Eisemann M, et al. Predictors of response to web-based cognitive behavioral therapy with high-intensity face-to-face therapist guidance for depression: a bayesian analysis. J Med Internet Res 2015 Sep 02;17(9):e197 [FREE Full text] [doi: 10.2196/jmir.4351] [Medline: 26333818]
50. Kay-Lambkin F, Baker A, Lewin T, Carr V. Acceptability of a clinician-assisted computerized psychological intervention for comorbid mental health and substance use problems: treatment adherence data from a randomized controlled trial. J Med Internet Res 2011 Jan 27;13(1):e11 [FREE Full text] [doi: 10.2196/jmir.1522] [Medline: 21273184]
51. 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]
52. Kolovos S, Kenter RM, Bosmans JE, Beekman AT, Cuijpers P, Kok RN, et al. Economic evaluation of internet-based problem-solving guided self-help treatment in comparison with enhanced usual care for depressed outpatients waiting for face-to-face treatment: A randomized controlled trial. J Affect Disord 2016 Aug;200:284-292. [doi: 10.1016/j.jad.2016.04.025] [Medline: 27155071]
53. Kok RN, van Straten A, Beekman AT, Cuijpers P. Short-term effectiveness of web-based guided self-help for phobic outpatients: randomized controlled trial. J Med Internet Res 2014 Sep 29;16(9):e226 [FREE Full text] [doi: 10.2196/jmir.3429] [Medline: 25266929]
54. Kenwright M, Liness S, Marks I. Reducing demands on clinicians by offering computer-aided self-help for phobia/panic. Feasibility study. Br J Psychiatry 2001 Nov;179:456-459 [FREE Full text] [Medline: 11689405]
55. Kiluk BD, Devore KA, Buck MB, Nich C, Frankforter TL, LaPaglia DM, et al. Randomized trial of computerized cognitive behavioral therapy for alcohol use disorders: efficacy as a virtual stand-alone and treatment add-on compared with standard outpatient treatment. Alcohol Clin Exp Res 2016 Sep;40(9):1991-2000. [doi: 10.1111/acer.13162] [Medline: 27488212]
56. Klein AA, Slaymaker VJ, Dugosh KL, McKay JR. Computerized continuing care support for alcohol and drug dependence: a preliminary analysis of usage and outcomes. J Subst Abuse Treat 2012 Jan;42(1):25-34. [doi: 10.1016/j.jsat.2011.07.002] [Medline: 21862275]

57. Kok RN, van SA, Beekman AT, Cuijpers P. Short-term effectiveness of web-based guided self-help for phobic outpatients: randomized controlled trial. J Med Internet Res 2014 Sep 29;16(9):e226 [FREE Full text] [doi: 10.2196/jmir.3429] [Medline: 25266929]
58. Kooistra LC, Ruwaard J, Wiersma J, 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. [doi: 10.1016/j.invent.2016.01.003]
59. Kordy H, Wolf M, Aulich K, Bürgy M, Hegerl U, Hüsing J, et al. Internet-delivered disease management for recurrent depression: a multicenter randomized controlled trial. Psychother Psychosom 2016;85(2):91-98. [doi: 10.1159/000441951] [Medline: 26808817]
60. Krieger T, Meyer B, Sude K, Urech A, Maercker A, Berger T. Evaluating an e-mental health program ("deprexis") as adjunctive treatment tool in psychotherapy for depression: design of a pragmatic randomized controlled trial. BMC Psychiatry 2014 Oct 08;14:285 [FREE Full text] [doi: 10.1186/s12888-014-0285-9] [Medline: 25298158]
61. Månsson KN, Skagius RE, Gervind E, Dahlin M, Andersson G. Development and initial evaluation of an internet-based support system for face-to-face cognitive behavior therapy: a proof of concept study. J Med Internet Res 2013 Dec 10;15(12):e280 [FREE Full text] [doi: 10.2196/jmir.3031] [Medline: 24326065]
62. Marks IM, Kenwright M, McDonough M, Whittaker M, Mataix-Cols D. Saving clinicians' time by delegating routine aspects of therapy to a computer: a randomized controlled trial in phobia/panic disorder. Psychol Med 2004 Jan;34(1):9-17. [Medline: 14971623]
63. Kim SJ, Marsch LA, Acosta MC, Guarino H, Aponte-Melendez Y. Can persons with a history of multiple addiction treatment episodes benefit from technology delivered behavior therapy? A moderating role of treatment history at baseline. Addict Behav 2016 Mar;54:18-23 [FREE Full text] [doi: 10.1016/j.addbeh.2015.11.009] [Medline: 26657820]
64. Nordmo M, Sinding A, Carlbring P, Andersson G, Havik O, Nordgreen T. Internet-delivered cognitive behavioural therapy with and without an initial face-to-face psychoeducation session for social anxiety disorder: A pilot randomized controlled trial. Internet Interv 2015 Nov;2(4):429-436. [doi: 10.1016/j.invent.2015.10.003]
65. Pier C, Austin DW, Klein B, Mitchell J, Schattner P, Ciechomski L, et al. A controlled trial of internet-based cognitive-behavioural therapy for panic disorder with face-to-face support from a general practitioner or email support from a psychologist. Ment Health Fam Med 2008 Mar;5(1):29-39 [FREE Full text] [Medline: 22477844]
66. Shandley K, Austin DW, Klein B, Pier C, Schattner P, Pierce D, et al. Therapist-assisted, internet-based treatment for panic disorder: can general practitioners achieve comparable patient outcomes to psychologists? J Med Internet Res 2008 May 19;10(2):e14 [FREE Full text] [doi: 10.2196/jmir.1033] [Medline: 18487138]
67. Reins JA, Ebert DD, Lehr D, Riper H, Cuijpers P, Berking M. Internet-based treatment of major depression for patients on a waiting list for inpatient psychotherapy: protocol for a multi-centre randomised controlled trial. BMC Psychiatry 2013 Nov 26;13:318 [FREE Full text] [doi: 10.1186/1471-244X-13-318] [Medline: 24279841]
68. Robertson L, Smith M, Castle D, Tannenbaum D. Using the internet to enhance the treatment of depression. Australas Psychiatry 2006 Dec;14(4):413-417. [doi: 10.1080/j.1440-1665.2006.02315.x] [Medline: 17116083]
69. 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 Dec 12;15:311 [FREE Full text] [doi: 10.1186/s12888-015-0697-1] [Medline: 26651478]
70. Sethi S, Campbell A, Ellis L. The use of computerized self-help packages to treat adolescent depression and anxiety. J Technol Hum Serv 2010 Aug 31;28(3):144-160. [doi: 10.1080/15228835.2010.508317]
71. Tannenbaum D, Spiranovic C. Severe, recurrent depression managed in a remote setting via the internet: an example of remote care using the healthsteps system. In: Taylor CB, editor. How to Practice Evidence-based Psychiatry: Basic Principles and Case Studies. Arlington, TX: American Psychiatric Publishing, Inc; 2011:275-286.
72. Van Voorhees BW, Ellis JM, Gollan JK, Bell CC, Stuart SS, Fogel J, et al. Development and process evaluation of a primary care internet-based intervention to prevent depression in emerging adults. Prim Care Companion J Clin Psychiatry 2007;9(5):346-355 [FREE Full text] [Medline: 17998953]
73. Whitfield G, Williams C, Hinshelwood R, Pashley A, Campsie L. The impact of a novel computerized CBT CD-Rom (overcoming depression) offered to patients referred to clinical psychology. Eur Psychiatry 2007 Mar;22:S248. [doi: 10.1016/j.eurpsy.2007.01.831]
74. Wilhelmsen M, Lillevoll K, Risør MB, Høifødt R, Johansen ML, Waterloo K, et al. Motivation to persist with internet-based cognitive behavioural treatment using blended care: a qualitative study. BMC Psychiatry 2013 Nov 07;13:296 [FREE Full text] [doi: 10.1186/1471-244X-13-296] [Medline: 24199672]
75. 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]
76. 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]

77. Carroll KM, Ball SA, Martino S, Nich C, Babuscio TA, Nuro KF, et al. Computer-assisted delivery of cognitive-behavioral therapy for addiction: a randomized trial of CBT4CBT. Am J Psychiatry 2008 Jul;165(7):881-888 [FREE Full text] [doi: 10.1176/appi.ajp.2008.07111835] [Medline: 18450927]
78. Marsch LA, Guarino H, Acosta M, Aponte-Melendez Y, Cleland C, Grabinski M, et al. Web-based behavioral treatment for substance use disorders as a partial replacement of standard methadone maintenance treatment. J Subst Abuse Treat 2014 Jan;46(1):43-51 [FREE Full text] [doi: 10.1016/j.jsat.2013.08.012] [Medline: 24060350]
79. Jacmon J, Malouff J, Taylor N. Treatment of major depression: effectiveness of cognitive behavior therapy with an internet course as a central component. EJAP 2010 Jan 03;5(2):8. [doi: 10.7790/ejap.v5i2.153]
80. Kenter RM, van SA, Hobbel SH, Smit F, Bosmans J, Beekman A, et al. Effectiveness and cost effectiveness of guided online treatment for patients with major depressive disorder on a waiting list for psychotherapy: study protocol of a randomized controlled trial. Trials 2013 Dec 01;14:412 [FREE Full text] [doi: 10.1186/1745-6215-14-412] [Medline: 24289099]
81. Murphy SM, Campbell AN, Ghitza UE, Kyle TL, Bailey GL, Nunes EC, et al. Cost-effectiveness of an internet-delivered treatment for substance abuse: Data from a multisite randomized controlled trial. Drug Alcohol Depend 2016 Apr 01;161:119-126 [FREE Full text] [doi: 10.1016/j.drugalcdep.2016.01.021] [Medline: 26880594]
82. Cahill J, Barkham M, Hardy G, Rees A, Shapiro DA, Stiles WB, et al. Outcomes of patients completing and not completing cognitive therapy for depression. Br J Clin Psychol 2003 Jun;42(Pt 2):133-143. [doi: 10.1348/014466503321903553] [Medline: 12828803]
83. Kenter RM, Cuijpers P, Beekman A, van Straten A. Effectiveness of a web-based guided self-help intervention for outpatients with a depressive disorder: short-term results from a randomized controlled trial. J Med Internet Res 2016 Mar 31;18(3):e80 [FREE Full text] [doi: 10.2196/jmir.4861] [Medline: 27032449]
84. 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 03;17(1):387 [FREE Full text] [doi: 10.1186/s13063-016-1511-1] [Medline: 27488181]
85. Lindhiem O, Bennett CB, Rosen D, Silk J. Mobile technology boosts the effectiveness of psychotherapy and behavioral interventions: a meta-analysis. Behav Modif 2015 Nov;39(6):785-804 [FREE Full text] [doi: 10.1177/0145445515595198] [Medline: 26187164]
86. Rose GL, Skelly JM, Badger GJ, Ferraro TA, Helzer JE. Efficacy of automated telephone continuing care following outpatient therapy for alcohol dependence. Addict Behav 2015 Feb;41:223-231 [FREE Full text] [doi: 10.1016/j.addbeh.2014.10.022] [Medline: 25452069]
87. Moore BA, Fazzino T, Barry DT, Fiellin DA, Cutter CJ, Schottenfeld RS, et al. The recovery line: a pilot trial of automated, telephone-based treatment for continued drug use in methadone maintenance. J Subst Abuse Treat 2013 Jul;45(1):63-69 [FREE Full text] [doi: 10.1016/j.jsat.2012.12.011] [Medline: 23375114]

Abbreviations

f2f: face-to-face
RCTs: randomized controlled trials

Edited by G Eysenbach; submitted 31.08.16; peer-reviewed by J Fogel, R van der Vaart, UD Reips; comments to author 25.10.16; revised version received 06.06.17; accepted 27.07.17; published 15.09.17.

Please cite as:
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
URL: http://www.jmir.org/2017/9/e306/
doi:10.2196/jmir.6588
PMID:28916506

Original Paper

Enhancing Seasonal Influenza Surveillance: Topic Analysis of Widely Used Medicinal Drugs Using Twitter Data

Ireneus Kagashe[1], MBA in ITM; Zhijun Yan[1,2], PhD; Imran Suheryani[3], Mphil

[1]School of Management and Economics, Beijing Institute of Technology, Beijing, China

[2]Sustainable Development Research Institute for Economy and Society of Beijing, Beijing, China

[3]School of Life Science, Department of Biomedical Engineering, Beijing Institute of Technology, Beijing, China

Corresponding Author:
Zhijun Yan, PhD
School of Management and Economics
Beijing Institute of Technology
Main Building
No. 5 South Zhongguancun Street, Haidian
Beijing, 100081
China
Phone: 86 10 68912845
Fax: 86 10 68912483
Email: yanzhijun@bit.edu.cn

Abstract

Background: Uptake of medicinal drugs (preventive or treatment) is among the approaches used to control disease outbreaks, and therefore, it is of vital importance to be aware of the counts or frequencies of most commonly used drugs and trending topics about these drugs from consumers for successful implementation of control measures. Traditional survey methods would have accomplished this study, but they are too costly in terms of resources needed, and they are subject to social desirability bias for topics discovery. Hence, there is a need to use alternative efficient means such as Twitter data and machine learning (ML) techniques.

Objective: Using Twitter data, the aim of the study was to (1) provide a methodological extension for efficiently extracting widely consumed drugs during seasonal influenza and (2) extract topics from the tweets of these drugs and to infer how the insights provided by these topics can enhance seasonal influenza surveillance.

Methods: From tweets collected during the 2012-13 flu season, we first identified tweets with mentions of drugs and then constructed an ML classifier using dependency words as features. The classifier was used to extract tweets that evidenced consumption of drugs, out of which we identified the mostly consumed drugs. Finally, we extracted trending topics from each of these widely used drugs' tweets using latent Dirichlet allocation (LDA).

Results: Our proposed classifier obtained an F_1 score of 0.82, which significantly outperformed the two benchmark classifiers (ie, $P<.001$ with the lexicon-based and $P=.048$ with the 1-gram term frequency [TF]). The classifier extracted 40,428 tweets that evidenced consumption of drugs out of 50,828 tweets with mentions of drugs. The most widely consumed drugs were influenza virus vaccines that had around 76.95% (31,111/40,428) share of the total; other notable drugs were Theraflu, DayQuil, NyQuil, vitamins, acetaminophen, and oseltamivir. The topics of each of these drugs exhibited common themes or experiences from people who have consumed these drugs. Among these were the enabling and deterrent factors to influenza drugs uptake, which are keys to mitigating the severity of seasonal influenza outbreaks.

Conclusions: The study results showed the feasibility of using tweets of widely consumed drugs to enhance seasonal influenza surveillance in lieu of the traditional or conventional surveillance approaches. Public health officials and other stakeholders can benefit from the findings of this study, especially in enhancing strategies for mitigating the severity of seasonal influenza outbreaks. The proposed methods can be extended to the outbreaks of other diseases.

(J Med Internet Res 2017;19(9):e315) doi:10.2196/jmir.7393

KEYWORDS

machine learning; Twitter messaging; social media; disease outbreaks; influenza; public health surveillance; natural language processing; influenza vaccines

Introduction

Background

Public health surveillance involves the systematic collection, management, analysis, and interpretation of health-related data, followed by the dissemination of these data to public health programs to enhance public health actions [1,2]. This process comes in handy during periods of severe health concerns such as disease outbreaks that need preventive and control interventions.

Among these disease outbreaks is influenza-like illness (ILI) or flu, which is a respiratory illness caused by viruses that can cause severe illness and even death for some people [3]. Common influenza types include seasonal, avian, swine, variant, and pandemic. Due to its recurring nature, we focused on seasonal influenza in this study.

To control disease outbreaks of seasonal influenza, one widely adopted measure is the proper use of medicinal drugs (ie, drugs which treat, prevent, or alleviate symptoms of diseases) [4]. When outbreaks occur, it is vitally important to know and analyze feedback data related to drugs that are widely consumed by people so that control measures to fight these outbreaks can be enhanced and implemented in the future.

Our work is built on the knowledge that seasonal influenza outbreaks affect a large number of people and are spread over large geographical areas. These facts pose challenges in gathering and analyzing useful feedback data related to the consumption of medicinal drugs by using traditional or conventional surveillance methods, which are limited by a small sample size, cost, and timeliness in reporting [5-7].

Therefore, this research study intended to extract relevant topics from tweets mentioning widely consumed drugs during seasonal influenza outbreaks and to use those topics to enhance seasonal influenza surveillance.

With the fast development of Web 2.0 technology, we considered using social media because of its ability to collect vast amounts of health-related data. Twitter, which was established in 2006, is among the most famous social media platforms and is currently the leading microblogging service for people to send and receive messages (tweets) of up to 140 characters. Twitter has a volume of 313 million monthly active users, 1 billion unique monthly visits, and 500 million tweets per day [8]. The communication forms in Twitter can be chats, conversations, news reporting, and information sharing [9]. People discuss their health conditions and statuses on Twitter [10], which makes it a new potential data source for health-related studies examining the prevalence of health issues, drug consumption, and health topics or categories [11-19].

To follow the use of influenza drugs in this study, we could have followed sales of over-the-counter and prescription drugs [5,20], or we could have used search engine queries and Web data [6,7,21,22]. However, prescription drug sales would not allow us to obtain topics or feedback from consumers, and access to these data is limited, even for researchers. As Web data and search engines also have limited access to researchers, we chose to use Twitter data.

However, selecting tweets that could be associated with actual uptake of drugs to determine which drugs are widely used and to derive insight from corresponding tweet topics is still a challenge.

Related Work

Medical Entity Estimates From Twitter

To obtain widely used drugs (or any other medical entity) from tweets, one needs to count the tweets with mentions of drugs (medical entities) and rank their frequencies. In previous studies, Twitter has been used to estimate the extent to which medical entities (drugs, diseases, symptoms) mentioned in tweets have been experienced or will be experienced by aggregating counts and then finding correlations with official surveillance data. For example, flu epidemics have been predicted in previous studies [11,23]. Correlations between flu or disease mentions in Twitter data and official flu case data have been found as well [12,18,19]. With regard to drugs, Twitter has been effectively leveraged to study prescription drug abuse [13], to track usage of illicit drugs [14], and to find a correlation between flu vaccine sentiment tweets and official vaccination rates according to geography [24].

The presence of medical entities in tweets does not necessarily signify that people have the illness, are using the pharmacological substances (drugs), or are exhibiting certain symptoms. Therefore, to obtain frequency of drug use from tweets, it is best to first identify which tweets indicate that tweeters have consumed the drugs mentioned in their tweets.

Several previous studies have paid attention to the identification of actual experiences of medical entities from other mentions. This includes work by Aslam et al [25] and Ji et al [26] who sought first to differentiate between *valid* and *invalid* tweets that expressed experiences with the flu before proceeding with further analyses. Weeg et al [27] found that the correlation between population disease prevalence and disease mentions in tweets increased from .113 to .208 ($P<.001$) when only disease name mentions that refer to actual diseases were taken into consideration. Using a somewhat similar approach, Alvaro et al [28] identified tweets that described firsthand experiences of prescription drugs, which were then used to gather evidence about adverse drug reactions.

Due to the large volumes of Twitter data, an effective way to separate the class of tweeters with experience with medical entities from the nonexperience class is by using machine learning (ML) techniques. However, researchers are still striving to improve the performance of the ML classifiers used for classifying tweets into those respective classes.

Topic Extraction From Twitter

To obtain insight from tweeters who have consumed these drugs, we sought to extract categories or topics from their respective tweets. Twitter data have been used in previous studies to extract categories or topics about health-related issues for various purposes. We have seen Twitter being used to investigate topics surrounding antibiotic prescription drugs [15] to find emerging trends for tweets related to electronic cigarettes [16], to detect topics at the peak of a disease related to official reports [29], and to find trending topics from preselected health-related keywords [17].

However, most of these studies conducted their topic analysis by classifying their tweets into predefined health categories [15,16] instead of obtaining these categories automatically. Those approaches miss opportunities to discover hidden and new topics. Additionally, the discovered topics were extracted from all the tweets that mentioned the medical entities [29]. However, only the tweets that expressed experiences with medical entities should be included in the topic extraction [17].

Influenza Drug Uptake Surveys

Uptake of drugs to treat or prevent influenza infections is an effective measure to mitigate the impact of influenza outbreaks. Consequently, there is a need to understand which factors contribute to the increase or decrease of drug uptake. With that information, public health agencies can take appropriate action against factors that may decrease drug uptake and promote factors that may increase drugs uptake.

Among the factors that deter uptake of influenza drugs (especially vaccines) are fear of needles, pain, and distress resulting from vaccination [30-32].

The logistics of vaccination processes that include the locations of vaccination centers and waiting time are among the factors that influence people to receive an influenza vaccine [32-34]. Some people tend to prefer traditional vaccination locations (hospitals, clinics, and doctor’s office) [32], whereas others prefer nontraditional locations (pharmacies, workplaces, and schools) [33,34]. Short line-ups and wait times help facilitate influenza vaccine uptake [34].

There have been concerns about the safety of the influenza drugs with respect to people with allergies [35] and pregnant women [36-39], which negatively affect the uptake of drugs.

Demographic factors, including age and level of education have been found to contribute to the uptake of influenza vaccines in which younger and less-educated adults are more hesitant to take drugs than other groups [40].

Public health agencies and pharmaceutical companies, as well as doctors or medical professionals [41,42] and parents [42] are also positive influences on drug uptake.

It has also been observed that some people prefer natural remedies to conventional drugs for influenza treatment or prevention [22]. However, natural remedies should be used to supplement and not replace tested influenza drugs.

The findings of these studies provide useful public health information for controlling the severity of seasonal influenza outbreaks through drug uptake. However, the methods for data collection used in most of these studies, that is, face-to-face interviews [30-32,37,41], written questionnaires [34,38], telephone surveys [33,42], secondary data databases [36], and Web page hits [22] have many drawbacks that may adversely affect the quality of the findings.

The cost for conducting these survey studies is one of the limitations. For example, a typical clinic-based survey (interview) for one participant in the United States costs US $23.51 compared with US $14.63 for a social media survey [43]. In another study [44], the cost of a telephone survey per sample (US $3.98) was higher compared with Web-based surveys (US $0.71). These high costs may hinder collection of data from large sample populations.

For public health agencies to take immediate and effective action, the findings of these studies should be released as soon as possible. However, these conventional methods are subject to delays with regard to data collection and may provide results later than they are needed to make informed decisions [45].

To employ effective strategies against seasonal influenza outbreaks, research results that reflect the actual situation in the population being studied should be used. However, because of the nature of the interactions between researchers and respondents in conventional or traditional methods, these studies are affected by social desirability bias in which respondents tend to provide responses that seem favorable [46]. This trend may lead to incorrect analysis results and ultimately cause public health agencies to fail to take appropriate action in locations where it is required.

Additionally, the size of study populations, which are often scattered in disparate geographical regions, can limit coverage in these studies. It is very costly to conduct studies with a population size that can be generalized to larger populations such as regions and countries.

In summary, existing approaches have two main limitations. First, the performance of ML classifiers used for identifying experience or personal tweets need improvements. These approaches mainly used n-gram bag-of-words or characters as ML features. These features suffer from the curse of dimensionality because the total dimension of each tweet’s text is equal to the vocabulary size, which can overfit the models. Additionally, these features do not consider semantic relations between words, which can result in poor performances in some cases.

Second, traditional surveys for finding insights about influenza drugs suffer from limitations associated with cost, timelessness in reporting, coverage, and bias.

Objectives

To achieve our research aim, we set the following objectives as our guidelines:

The first objective was to provide a methodological extension for efficiently extracting more widely used drugs during seasonal influenza using tweets with evidence for consumption of drugs. We focused on providing an improved ML classifier that could identify tweets indicating uptake of drugs from others. We hypothesized that using a dependency words structure (introduced in the Methods section) of the tweets as our features would improve performance in classification.

Our second objective was to extract topics from tweets mentioning each of the widely consumed drugs and to infer how the insights provided by these topics can enhance seasonal influenza surveillance in lieu of relying on traditional surveillance. From this perspective, we focused on automatically (without predefined categories) finding fine-tuned topics (extracted separately from tweets about each drug). We hypothesized that topics of widely consumed drugs could be used to enhance surveillance of seasonal influenza. Additionally, the tweets from these topics were the ones that signified actual consumption of drugs.

The contribution or significance of the research is two-fold in the following ways:

First, we proposed a new enhanced classification method to identify whether a tweet indicates someone has consumed or intended to consume a drug or not. This enhanced classification method guarantees results that reflect the actual situation in the population being studied when searching for widely used drugs and in subsequent analyses (topic extraction).

Second, a topic extraction–based method was applied to analyze the hot topics in tweets that can help public health stakeholders enhance seasonal influenza surveillance and intervention measures in terms of drug administration and consumption. Instead of using traditional or conventional survey methods, the topic extraction–based method can directly extract topics from tweets of people who have consumed these drugs. This sort of approach guarantees cost-effective, fast, and high coverage results for public health stakeholders to take action if needed.

To the best of our knowledge, this paper is the first to use a topic extraction–based method to retrieve insights about seasonal influenza drugs using Twitter data. The retrieved topics can highlight enabling and deterrent factors to drug uptake.

We evaluated our ML classifier by comparing it with a 1-gram term frequency (TF)–based classifier and a clue (keyword or lexicon)-based classifier. These benchmarks were chosen because they have been extensively used for separating tweets of individuals with actual experience (personal or valid) from tweets made by others [25,26,28,47-50]. For the extracted topics, we examined how the findings support the results of previous studies that used conventional or traditional surveillance methods.

Methods

Overview

The overview of the proposed method is summarized in Figure 1. The tweets are first preprocessed and filtered to obtain tweets with drug mentions only. Then, a random sample of these tweets is manually annotated as relevant or irrelevant with respect to actual drug uptake evidence. Next, features to be used by our classifier are generated. The classifier is trained, evaluated, and applied to the whole dataset to identify tweets indicative of the consumption of drugs (relevant tweets). From these tweets, we derive (1) a ranked list of widely used drugs and (2) topics from tweets mentioning each of these widely used drugs. A detailed description and implementation of these steps are provided in the following sections.

Figure 1. An overview of the proposed method.

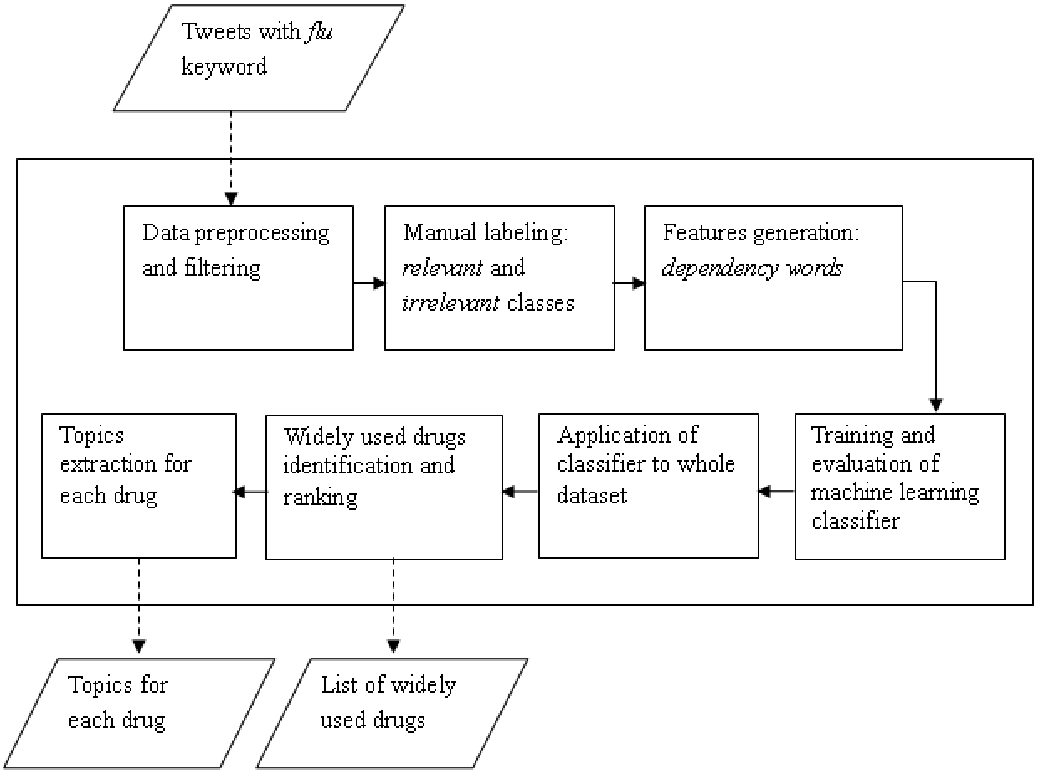


Table 1. Samples of relevant and irrelevant tweets with mentions of drugs (drugs italicized).

Category	Tweet text
Relevant	*Nyquil* cold&ampFlu thanks for my life back!
	Got my *flu shot*. Ughhh.
	I got the *flu vaccine* that means I am gunna get sicky.
Irrelevant	We have *flu shots* In! Make Your Appointment Today 642-###
	He got a song about some damn *Thera flu* lol
	So I either have the flu or a mild case west nile virus! *FML*!!

Data

We used historical Twitter data for the 2012-2013 influenza season collected from August 31, 2012 to March 04, 2013 within 30 cities in the United States using the Twitter application programming interface (API) and the keyword *flu*. This approach ensured that the analyzed drugs were related to seasonal flu outbreaks. Only tweets written in the English language were considered, and in total there were 459,043 tweets included in the analysis.

Post collection filtering was done to remove tweets that contained the keyword *flu* but were unrelated to the flu disease. For example, some tweets contained the phrase *stomach flu*, which implies an intestinal infection. Additionally, tweets that had the keyword *flu* as part of the Twitter username or handle or Twitter hashtag (such as @this_is_flu or #flu_camp, respectively) were also removed from our dataset. Because the majority of retweets and tweets with URLs do not usually express actual experiences of health-related issues [25], they were also removed from the dataset.

To identify the tweets that mentioned drugs, the unified medical language system (UMLS)-MetaMap [51] developed at the US National Library of Medicine was used. The semantic type of each token or word in a tweet was first identified by the UMLS-MetaMap. If the word was classified into one of the following three UMLS semantic types: immunological factor, clinical drug, or pharmacological substance (which are the semantic types for preventive or treatment drugs), then the corresponding tweet was kept for subsequent analyses.

Manual Annotation of Tweets

Since we wanted to train an ML classifier that could identify tweets indicating uptake of drugs, we first manually labeled a random sample of the tweets in two classes, including the *relevant* class if a tweet indicated actual consumption of drugs and the *irrelevant* class for other topics. These manually labeled data were later used to train and test the classifier. Table 1 has some examples of relevant and irrelevant tweet labeling.

Two annotators were tasked with categorizing 5000 tweets into the two classes (relevant and irrelevant). The interrater agreement between the 2 annotators measured using Cohen kappa [52] was .91. A third annotator was recruited to decide on the disagreements between the 2 annotators through majority votes.

Machine Learning Classifier Features

The proposed classifier used dependency words as features that were obtained as follows: We parsed each tweet to obtain the dependency structure of words, and then we identified which words in the tweet had dependency relations with a drug mentioned in the tweet; we only used these words as our features. Our approach was peculiar in the sense that we did not use the dependency grammar categories (eg, nsbuj, dobj, and conj) of the words in the tweets as features. We named this feature *dependency words*.

TF-inverse document frequency (IDF) basically represents each text document (in this case, a tweet's dependency words) as a vector of terms or words with each component of the vector corresponding to each term in the corpus and has a weight or count associated with it [52]. The TF part measures how frequently a term occurs in a document, whereas the IDF part weights the most frequent terms while scaling up the rare ones. For corpus terms that are not in the document, their weights are zero. The terms can have any n-gram (n=1, 2...) size. However, Cole-Lewis et al [16] determined that using 1-gram instead of other n-grams to classify tweets into relevant and irrelevant categories led to the best performance. Since we intended to implement our model with a support vector machine (SVM) algorithm that also scales the terms, we used TF only. Thus, a 1-gram TF of dependency words was adopted in this study to categorize tweets into relevant or irrelevant types.

The Stanford parser (Stanford University) [53] was used to find dependency words, whereas during TF feature construction, we used the Stanford CoreNLP [54] to reduce the dependency words into their lemma forms.

The following example shows the dependency words extracted from a sentence in a tweet:

> Original sentence: *Just got my Nyquil, ready for bed*.
>
> Results: *my*, *got*, and *ready*

From the example, the drug mentioned in the tweet is Nyquil, which has grammatical dependency relations with the words *my*, *got*, and *ready*.

Classifier Training, Evaluation, and Application

We used the SVM with a linear kernel function as the ML algorithm, which was implemented using LIBSVM from the National Taiwan University [55]. To obtain a trained SVM model that can provide optimum classification results, there are two parameters that must be tuned, namely C and gamma. We applied a grid search method on C and gamma using 10-fold cross validation to test several pairs of C and gamma, and the pair that provided the best 10-fold cross-validation performance for the training set was chosen. With these parameters, the classifier was then trained with 3400 labeled data, and its performance was evaluated using 1600 test data in terms of precision, recall, and F_1 score. To ascertain the classifier's ability to avoid overfitting, the classifier was tested by evaluating its performance on the test dataset that was not used for training. This test dataset was assumed to approximate the typical unseen data the classifier would encounter in the future.

Our training set had a total of 701 features (dependency words). We tested whether we could achieve the same or better performance with fewer features through feature selection (20%, 40%, 60%, and 80%), but we could not get better performance with less than 100% of the features. Therefore, all of the features were used.

To evaluate the performance of our proposed ML classifier, we compared it with two benchmark classifiers, including a 1-gram TF-based classifier used in prior studies [25,28,47-49] and a lexicon (keyword or clue)-based classifier as used in another previous study [26]. The 1-gram TF classifier was constructed by using all of the words in each tweet after removing all of the named entities (eg, person, location, organization, numerical, and temporal) and stop words, which are considered to have little or no contribution to document classification. The lexicon-based classifier was constructed by considering the presence of subjectivity words, news keywords, and profanity words. A tweet was considered to indicate drug experience (relevant) if it contained at least 3 strong subjective words and 3 weak subjective words. In contrast, a tweet was considered to be irrelevant if it contained news keywords, and there were no profane words.

The optimum-trained model was then applied to the whole corpus of tweets to find all relevant tweets, that is, tweets indicating consumption of drugs. We counted the drugs mentioned in these tweets and ranked them to identify widely consumed drugs.

Topic Extraction for Tweets Mentioning Drugs

Topic models are based on a concept that documents are mixtures of topics in which a topic is a probability distribution for words [56] . In other words, a topic refers to a group of words that frequently occur together and can form meaningful and interpretable themes. Latent Dirichlet allocation (LDA) is one of the simplest topic models and is widely used in Web text mining.

LDA is a generative model for topic modeling that can automatically discover hidden topics from a collection of text documents represented as a bag-of-words [57]. Each document is regarded as a mixture of several topics, and a topic is a distribution of words. For a collection of text documents, the LDA model can generate a certain number of topics. By understanding the topic distributions among text documents and the word distributions among topics, unknown or hidden information in the text can be retrieved.

The goal of topic modeling was to separately discover hidden topics from a collection of relevant tweets for each widely used drug. For the tweets of each drug, we applied the LDA topic modeling method to find topics. The LDA method was implemented by MALLET (University of Massachusetts Amherst) [58].

The number of topics retrieved for tweets about each drug was varied using an optimum topic number test as suggested by a previous method [59]. We applied the LDA topic model to the documents (tweets) with a randomly specified number of topics and observed the per-document topic distributions results. If the per-document topic distributions of all documents were dominated by a few topics, then the number of topics was increased and vice versa, until an optimal balance was found. The only parameter or value we specified was the number of topics. Other LDA parameters were tuned automatically from their default values (alpha=5.0 and beta=.01) based on the specified number of topics and based on the number of words in the tweets. Detailed information on our implementation of the LDA-based topic model for tweets of widely consumed drugs is provided in Multimedia Appendix 1.

The results of this LDA topic model included per-document topic distributions, a set of topics, and word-to-topic assignments for each word in the corpus. However, the data that were most interesting to us were the extracted topics.

Results

Relevant Tweets With Drug Mentions

After the initial preprocessing and filtering of tweets to remove retweets, as well tweets with URLs and tweets with the keyword *flu* that did not imply the flu disease, we had 220,375 tweets remaining. The number of tweets with drug mentions was 50,828 after filtering with MetaMap software. The number of relevant tweets that indicated actual uptake of drugs after applying our classifier was 40,428. These 40,428 tweets produced 6232-word vocabulary size (ie, number of dependency words).

Classifier Performance Evaluation

The performance of our SVM classifier, when applied to the test set measured in terms of precision, recall, and the F_1 score is shown in Table 2. The gamma and C parameter values for the SVM classifier that provide the optimum performance were .008 and 8, respectively, after a thorough grid search. Our proposed model, which was 1-gram TF of dependency words, outperformed the model constructed using 1-gram TF and the lexicon-based model as depicted in Table 2.

The test dataset was divided into 30 parts, and a statistical significance test of our classifier over the two benchmark classifiers was conducted to check whether the proposed method outperformed the two baseline methods significantly. The results showed that the differences in the F_1 score between our classifier and the two benchmarks were statistically significant (ie, P<.001 with the lexicon-based model and P=.048 with the 1-gram TF model).

Table 2. Performance of the classifiers using lexicon-based, 1-gram term frequency (TF), and dependency word features.

Classifier features	Precision	Recall	F_1 score
Lexicon-based (benchmark 1)	0.52	0.91	0.66
1-gram TF[a] (benchmark 2)	0.73	0.88	0.79
Dependency words (our approach)	0.77	0.90	0.82

[a] TF: term frequency.

Table 3. Widely used drugs retrieved from relevant tweets (N=40,428).

Drugs	Tweets count, n (%)
Influenza virus vaccines	31,111 (76.95)
Theraflu	1267 (3.13)
Vitamins	439 (1.09)
NyQuil	354 (0.88)
Acetaminophen	270 (0.67)
Oseltamivir	162 (0.40)
DayQuil	75 (0.20)

Table 4. Relevant topics retrieved from tweets mentioning influenza virus vaccines (only interpretable topic compositions are listed).

Topic number[a]	Topic compositions
1	mom, today, needles, doctor, nurse, gave, give, told, dad, baby, big, shot, giving, making, wanted, lady, sh*t, f*ck
2	reaction, hoping, sick, kids, allergic, eggs, made, egg, chicken, allergy, medicine, tea
3	flu, season, people, year, virus, immune, system, shot, epidemic, strain, stay, healthy, spreading, protect, remember
4	influenza, pregnant, risk, vaccination, national, recommend, women, immunity, free, safe
5	arm, sore, today, hurts, hurt, yesterday, damn, left, feels, side, bad, feel, stupid, throat, pain, ouch, killing, feeling, hurting
6	waiting, cvs, line, walgreens, pharmacy, free, wait, office, give, long, gave, spray, clinic, nasal, giving, people
7	sick, hate, shots, f*ck, sh*t, flu, needles, damn, today, nervous
8	work, office, day, today, free, morning, tomorrow, doctors, doctor, shot, good, school

[a] Corresponding interpretations of these topics are in Table 5.

Widely Used Drugs

Table 3 shows the counts for the selected widely used drugs and their percentages. We only considered drugs with tweet count percentages that were at least 0.20%. We excluded some *drugs* that the MetaMap software identified as drugs, but the tweets did not imply drug use. We did a manual analysis of these *drugs* by going through some of the tweets they came from. For example, *RID* in many of the tweets was used to mean *to clear or to free*, but it was identified as a drug because it could also mean pyrethrins or piperonyl in other contexts. This outcome occurred because the syntactic structure of tweets with these *drugs* was similar to tweets with mentions of true drugs, which made it difficult for the classifier to detect this ambiguity automatically.

Since we wanted to obtain frequencies or counts of drug uptake only, we did not include water and disinfectant products in the table, which were also identified as widely used drugs.

Although some drugs were identified separately by the MetaMap software, they are actually the same. Therefore, we grouped these drugs into one category for the sake of simplicity. For example, flu shots, flu vaccines, flu vaccine, vaccines, and vaccine were combined and presented as one category or class called influenza virus vaccines, and vitamins and vitamin C were grouped into vitamins.

Topics for Drugs-Mentioning Tweets

The optimum number of topics retrieved for each *relevant* drug was as follows: Influenza virus vaccines (10), Theraflu (5), DayQuil and NyQuil (5), vitamins (5), acetaminophen (3), and oseltamivir (3).

Table 4 shows topics of tweets with influenza virus vaccine mentions. Multimedia Appendix 2 has the full results of topic extraction for all drugs, which includes the number of topics, topic compositions, and LDA parameters.

In Table 5, we presented interpretations of the topics (clusters of frequently occurring words) for the drugs listed in Table 3. These interpretations express the meanings that the topics convey.

Table 5. Interpretations of topics retrieved for each drug.

Drugs	Topic interpretations
Influenza virus vaccines[a]	Vaccination proponents, vaccination allergic reaction, vaccination reminders, vaccination pregnancy risk, vaccination pain and distress, vaccination queues concerns, vaccination fear, vaccination places
Theraflu	Natural flu remedies uptake (chicken soup, hot drinks)
DayQuil or NyQuil	Drug uptake time (morning, night)
Vitamins	Flu preparedness through vitamins intake
Acetaminophen	Symptoms; Natural flu remedies uptake (soup, tea, orange juice)
Oseltamivir	Prescription of drug

[a] Correspond to topic numbers in Table 4.

Discussion

Principal Findings and Comparison With Previous Studies

The Classification Method

When compared with the two benchmarks, our proposed method showed performance (F_1 score) improvement for classifying tweets with mentions of drugs into *relevant* and *irrelevant* categories. This improvement was significant compared with both the 1-gram TF classifier (*P*=.048) and the lexicon-based classifier (*P*<.001).

The ML classification intention was designed to find which tweets indicate that a tweeter has consumed or intends to consume the drugs mentioned in their tweets. We used dependency parsing to first find which words in the tweets were involved in the binary grammatical relations with the drugs mentioned in the tweets, and then by using these words as our ML classifier features, we achieved better performance. This result occurred because these words (features) are closely related or aligned to the drugs and serve as better distinguishing features than using all words (1-gram TF) or even had much better results than using lexicon-based methods. Additionally, as dependency words were used only as features, the approach is not affected by dimensionality, and there is no possibility of model overfitting.

The two benchmarks were chosen because of the fact that they have been extensively used to identify *valid or relevant or personal* and *invalid or irrelevant or nonpersonal* tweets regarding flu experience, which is similar to the procedures used in prior studies [25,26,28,47-50].

Widely Used Drugs and Corresponding Topics

In this study, we defined *relevant* tweets as the ones that were composed by people who express their or another person's consumption of drugs. Out of 50,828 tweets that mentioned drugs, 40,428 (79.54%) were relevant tweets. This result implied that people discuss health-related conditions and actions on Twitter for both themselves and other people. This outcome is consistent with the work of Yin et al [10] who investigated whether Twitter users disclose health statuses (either their status or the status of other people).

The research estimated the proportions of different drugs used during the flu season. The results showed that widely used drugs included influenza virus vaccines 76.95% (31,111/40,428). Previous studies also investigated the uptake of influenza drugs, especially vaccines [60,61]. However, these studies mostly relied on surveys or interviews that were prone to social desirability bias, delays, and were very costly. Our approach leveraged tweets to obtain these proportions quickly and at a low cost, which can efficiently reflect the actual drug uptake.

Furthermore, in this research, we found that the topics of tweets mentioning the drugs varied depending on the types of drugs that were discussed. These topics could not be easily and accurately found using normal search queries [15] or traditional survey methods because of their limited scope. Combining mentions of widely used drugs together with trending topics can be beneficial to various public health stakeholders for controlling flu epidemics.

Regarding influenza virus vaccines, we observed that during flu seasons, people who were vaccinated tended to remind others to do the same to protect themselves from the flu and stay healthy (Table 4, topic 3: flu, season, people, year, virus, immune, system, shot, epidemic, strain, stay, healthy, spreading, protect, remember). This is a positive sign because lack of reminders has been found to be one of the barriers to vaccination among adults [62]. The following tweet shows how people who get vaccinated urge others to do so:

> *Guys protect yourself this flu season. Just got my mandatory flu shot. #HealthyPeople*

Additionally, apart from medical personnel, we noticed that people got their flu vaccines because of the persuasion or presence of their parents (Table 4, topic 1: mom, today, doctor, nurse, gave, give, told, dad, baby, big, shot, giving, making, wanted, lady, sh*t, f*ck). The presence of *curse words* and *needle* indicates pain resulting from vaccine injections, whereas the words *doctors*, *nurse*, *moms*, and *dads* indicate that vaccines were given or offered in the presence of these people or these people influenced the vaccination process. This result provides information on which cohort to target for sensitization campaigns to have a more positive impact on flu medication uptake, especially preventive drugs such as flu vaccines or shots. According to Giese et al [63], individuals are likely to be vaccinated if their doctors or medical staff have recommended the vaccine. Our topics analysis has discovered another potential group (parents) that can also facilitate convincing more people to get vaccinated because they were involved in the vaccination process. We expected parents to make vaccination decisions for young children only [64,65]. However, as Twitter users are mainly adults, it appears that parents can influence adults too:

> *...my moms making me get the flu shot):*
>
> *Moms making me get the flu shot. I f****ng hate needles.*
>
> *Idk my dad wants me to get a flu shot he knows I hate needles.*

Additionally, the extracted topics indicated that people got or had intended to get their vaccines at places where they were spending most of their time during the day and where the drugs were offered or given for free (Table 4, topic 8: work, office, day, today, free, morning, tomorrow, doctors, doctor, shot, good, school). The presence of topic words such as *work*, *office*, *school*, and *free* means that vaccines were given at those places (mostly for free). The finding can imply that offering drugs at these places can be an effective decision because of the peer pressure effect, rather than offering these drugs at health facilities only. This result correlated with the finding that people who receive vaccines are individuals who are in contact with many others during their daily activities [66]:

> *My office is giving out free flu shots I think I will go, last year my insurance did not cover my flu shot.*

The topic analysis also showed that people who got or had intended to get their vaccines were concerned with waiting for a long time in vaccination queues (Table 4, topic 6: waiting, cvs, line, walgreens, pharmacy, free, wait, office, give, long, gave, spray, clinic, nasal, giving, people). The words *line*, *long*, and *waiting* meant that waiting lines were long, or people waited for a long time in lines. The words *cvs*, *walgreens*, *pharmacy*, *office*, and *clinic* are vaccination locations. These results imply that time spent waiting on long lines at vaccination centers was a concern for many people, and measures need to be taken to address that concern:

> *2 hours in line at cvs for a flu shot #sickofthissh*t.*

Additionally, the aftermath of pain and distress from vaccines were among the concerns of many people (Table 4, topic 5: arm, sore, today, hurts, hurt, yesterday, damn, left, feels, side, bad, feel, stupid, throat, pain, ouch, killing, feeling, hurting). The presence of words such as *hurts*, *hurt*, *hurting*, *sore*, *damn*, *ouch*, and *pain* indicates pain, whereas the words *arm* and *left* indicate the body part affected by the pain. This finding can help health administration to design better strategies for offering these drugs and to avoid discouraging intended recipients in the future. For example, the 5P (procedural, physical, pharmacologic, psychological, and process) pain management intervention strategies [67,68] can be employed to ease pain and distress from drugs. The following tweets indicate the pain and distress users tend to associate with vaccine use:

> *I got the flu shot today and my arm STILL HURTS I'm all sore.*
>
> *This FLU SHOT after effect is KILLING ME. My arm so sore and in pain OMG.*

Vaccination needle fear was also among the themes that emerged from the tweeters (Table 4, topic 7: sick, hate, shots, f*ck, sh*t, flu, needles, damn, today, nervous). This means sensitization needs to be conducted to ensure that more people overcome their fear and partake in vaccination. The following tweet indicates the fear of needles:

> *Getting a flu shot. So nervous. Hate shots.*

Tweeters with egg or chicken allergies raised concerns about whether the vaccines would cause an allergic reaction (Table 4, topic 2: reaction, hoping, sick, kids, allergic, eggs, made, egg, chicken, allergy, medicine, tea). This outcome implies proper education should be given to avoid scaring away people with egg or chicken allergies or other allergies:

> *@Twitterhandle That's good to know. Both me and my little girl have egg allergy and got no ill effects from the flu shots.*

There were also themes showing concerns about the risk of vaccination to pregnant women (Table 4, topic 4: influenza, pregnant, risk, vaccination, national, recommend, women, immunity, free, safe). This finding means people were concerned with the efficacy and safety of the vaccines, which is similar to the findings of another study [69] and suggests there should be strategies to improve vaccination uptake among pregnant women.

> *1st getting my flu shot. Is it safe to get a flu shot during pregnancy? Yes. Better safe than sorry.*

Some topics for Theraflu and acetaminophen were related to how people who consumed these drugs also tended to use natural or home flu remedies (Multimedia Appendix 2, table MA3, topic 1: soup, chicken, noodle, juice, orange, care, flu, work, sleepy, warm, easier, food, drinkin, spicy, vodka) and (Multimedia Appendix 2, table MA5: fever, throat, juice, recommend, drowsy, soup, temp, symptom, strep, spray, orange, tea). However, there is a need to provide awareness to people so that they use these home remedies only as early interventions or in conjunction with clinically prescribed drugs for effective treatment of the flu [70]. The following tweets indicate that some people use drugs such as Theraflu and acetaminophen in conjunction with home remedies:

> *Chicken noodle soup and thera flu for dinner...mmm.*
>
> *@Twitterhandle drink some green tea and chicken soup and take theraflu and get rest#lets earn two.*

For the other drugs (except for influenza vaccines, Theraflu, and acetaminophen), the observable interpretable topics were related to preparedness against flu, symptoms, and prescriptions, as well as the use of the drugs to prevent or combat the flu.

Overall, in this research, we went the extra mile by ensuring that the topics about drugs reflected actual experiences of medical entities or drugs and that not all mentions of drugs were included, as they were in prior studies [17,29]. Additionally, these topics were supported by findings from several studies that are related to influenza surveillance but used traditional survey methods.

Limitations and Future Work

The study succeeded in improving the performance of the classifier when compared with the benchmarks and managed to find important themes about the consumed drugs. However, the study was limited by challenges such as misspellings, abbreviations, and slang language, which are common issues facing researchers attempting to analyze social media text. As most tweets were composed by individuals with no medical backgrounds, the tweets had lots of misspellings when mentioning drugs. Additionally, because of space restrictions and the informal nature of Twitter conversations, people often opted to use abbreviations and slang when exchanging messages. In this study, we automatically extracted treatment drugs using MetaMap software, which recognized only standard medical names and left out some tweets with drug mentions. This might have caused the number of tweets with a specific drug to be reduced and made it hard to find all of the meaningful topics. To overcome these challenges, future research studies should consider using MetaMap software combined with appropriate misspelling identification and correction methods. Additionally, abbreviations and slang should be expanded to their full forms and interpreted using appropriate lexica.

Additionally, we did not explore correlations of the uptake of the drugs between tweets and official statistics. We also did not conduct analyses in smaller, more specific geographical locations (such as cities). Instead, we considered the country as a whole. Future research studies can conduct analyses at the city level and find correlations between results from tweets and official data.

Conclusions

As the number of users and the sharing of health information on Twitter increase, Twitter has turned out to be a potential data source for public health research. It can help to investigate the uptake of various drugs and perceptions of users toward those drugs during seasonal flu outbreaks. Analyzing these massive datasets requires efficient methods that can identify emerging trends in the uptake and administration of drugs. Using ML techniques, this study proposed a methodological extension for efficiently extracting more widely used drugs during seasonal influenza using tweets confirming the consumption of drugs during seasonal flu outbreaks from a pool of tweets with mentions of drugs. Emerging topics of each confirmed tweet about drugs were extracted, which provided hidden information conveyed by people who have consumed these drugs. Mainly, these insights included encouraging or discouraging factors for influenza drug uptake. This information was obtained automatically rather than being obtained by conventional methods that have many shortcomings. Therefore, public health entities and other stakeholders can make full use of this efficiently obtained information to devise efficient and effective strategies for influenza epidemic surveillance.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (Award #:71572013, 71272057, 71521002) and Joint Development Program of Beijing Municipal Commission of Education. Findings, opinions, conclusions, or recommendations expressed in this research are those of the authors and do not necessarily reflect the views of the National Natural Science Foundation of China and Beijing Municipal Commission of Education.

The authors would like to extend their gratitude to Professor Ming-Hsiang Tsou from Geography Department, San Diego State University for sharing raw Twitter data used in this study.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Latent Dirichlet allocation (LDA) topic model and implementation.

[PDF File (Adobe PDF File), 547KB - jmir_v19i9e315_app1.pdf]

Multimedia Appendix 2

Topics for tweets with drug mentions.

[PDF File (Adobe PDF File), 343KB - jmir_v19i9e315_app2.pdf]

References

1. Porta M, Greenland S, Silva I, Last JM. A dictionary of epidemiology. New York: Oxford University Press; 2014.
2. Lee LM, Teutsch SM, Thacker SB, St Louis ME. Principles & practice of public health surveillance, 3rd edition. Oxford: Oxford University Press; 2010.
3. CDC. 2017. Seasonal influenza (flu): about flu URL: https://www.cdc.gov/flu/about/index.html [accessed 2017-08-19] [WebCite Cache ID 6sp9FqLKg]
4. CDC. 2016. CDC says “Take 3” actions to fight the flu URL: http://www.cdc.gov/flu/protect/preventing.htm [accessed 2016-11-16] [WebCite Cache ID 6m2KJkwd9]

5. Sugawara T, Ohkusa Y, Ibuka Y, Kawanohara H, Taniguchi K, Okabe N. Real-time prescription surveillance and its application to monitoring seasonal influenza activity in Japan. J Med Internet Res 2012 Jan 16;14(1):e14 [FREE Full text] [doi: 10.2196/jmir.1881] [Medline: 22249906]
6. Seo D, Jo M, Sohn CH, Shin S, Lee J, Yu M, et al. Cumulative query method for influenza surveillance using search engine data. J Med Internet Res 2014 Dec 16;16(12):e289 [FREE Full text] [doi: 10.2196/jmir.3680] [Medline: 25517353]
7. Sharpe JD, Hopkins RS, Cook RL, Striley CW. Evaluating google, twitter, and wikipedia as tools for influenza surveillance using bayesian change point analysis: a comparative analysis. JMIR Public Health Surveill 2016 Oct 20;2(2):e161 [FREE Full text] [doi: 10.2196/publichealth.5901] [Medline: 27765731]
8. Twitter. 2016. Twitter usage/company facts URL: https://about.twitter.com/company [accessed 2016-11-15] [WebCite Cache ID 6m2MLDMe2]
9. Java A, Song X, Finin T, Tseng B. Why we twitter: understanding microblogging usage and communities. In: Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 Workshop on Web mining and Social Network Analysis.: ACM; 2007 Presented at: WebKDD/SNA-KDD '07; August 12, 2017; San Jose, CA p. 56-65. [doi: 10.1145/1348549.1348556]
10. Yin Z, Fabbri D, Rosenbloom ST, Malin B. A scalable framework to detect personal health mentions on twitter. J Med Internet Res 2015;17(6):e138 [FREE Full text] [doi: 10.2196/jmir.4305] [Medline: 26048075]
11. Aramaki E, Maskawa S, Morita M. Twitter catches the flu: detecting influenza epidemics using twitter. In: Proceedings of the Conference on Empirical Methods in Natural Language Processing. 2011 Presented at: EMNLP 2011; July 27-31, 2011; Edinburgh p. 1568-1576.
12. Paul M, Dredze M. You are what you tweet: analyzing twitter for public health. In: Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media.: AAAI Press; 2011 Presented at: ICWSM-11; July 17–21, 2011; Barcelona p. 265-272 URL: https://www.aaai.org/ocs/index.php/ICWSM/ICWSM11/paper/view/2880/3264
13. Hanson CL, Burton SH, Giraud-Carrier C, West JH, Barnes MD, Hansen B. Tweaking and tweeting: exploring Twitter for nonmedical use of a psychostimulant drug (Adderall) among college students. J Med Internet Res 2013;15(4):e62 [FREE Full text] [doi: 10.2196/jmir.2503] [Medline: 23594933]
14. Daniulaityte R, Nahhas RW, Wijeratne S, Carlson RG, Lamy FR, Martins SS, et al. “Time for dabs”: analyzing twitter data on marijuana concentrates across the U.S. Drug Alcohol Depend 2015 Oct 1;155:307-311. [doi: 10.1016/j.drugalcdep.2015.07.1199] [Medline: 26338481]
15. Kendra RL, Karki S, Eickholt JL, Gandy L. Characterizing the discussion of antibiotics in the twittersphere: what is the bigger picture? J Med Internet Res 2015;17(6):e154 [FREE Full text] [doi: 10.2196/jmir.4220] [Medline: 26091775]
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;17(8):e208 [FREE Full text] [doi: 10.2196/jmir.4392] [Medline: 26307512]
17. Yepes A, MacKinlay A, Han B. Investigating public health surveillance using twitter. In: Proceedings of the 2015 Workshop on Biomedical Natural Language Processing (BioNLP 2015). 2015 Presented at: ACL-IJCNLP 2015; July 30, 2015; Beijing p. 164-170. [doi: 10.18653/v1/W15-3821]
18. Gesualdo F, Stilo G, Agricola E, Gonfiantini MV, Pandolfi E, Velardi P, et al. Influenza-like illness surveillance on twitter through automated learning of naïve language. PLoS One 2013;8(12):e82489 [FREE Full text] [doi: 10.1371/journal.pone.0082489] [Medline: 24324799]
19. Broniatowski DA, Paul MJ, Dredze M. National and local influenza surveillance through twitter: an analysis of the 2012-2013 influenza epidemic. PLoS One 2013;8(12):e83672 [FREE Full text] [doi: 10.1371/journal.pone.0083672] [Medline: 24349542]
20. Pivette M, Mueller JE, Crépey P, Bar-Hen A. Drug sales data analysis for outbreak detection of infectious diseases: a systematic literature review. BMC Infect Dis 2014 Nov 18;14:604 [FREE Full text] [doi: 10.1186/s12879-014-0604-2] [Medline: 25403237]
21. Woo H, Cho Y, Shim E, Lee J, Lee C, Kim SH. Estimating influenza outbreaks using both search engine query data and social media data in South Korea. J Med Internet Res 2016;18(7):e177 [FREE Full text] [doi: 10.2196/jmir.4955] [Medline: 27377323]
22. Hill S, Mao J, Ungar L, Hennessy S, Leonard CE, Holmes J. Natural supplements for H1N1 influenza: retrospective observational infodemiology study of information and search activity on the Internet. J Med Internet Res 2011 May 10;13(2):e36 [FREE Full text] [doi: 10.2196/jmir.1722] [Medline: 21558062]
23. Signorini A, Segre AM, Polgreen PM. The use of twitter to track levels of disease activity and public concern in the U.S. during the influenza A H1N1 pandemic. PLoS One 2011;6(5):e19467 [FREE Full text] [doi: 10.1371/journal.pone.0019467] [Medline: 21573238]
24. Salathé M, Khandelwal S. Assessing vaccination sentiments with online social media: implications for infectious disease dynamics and control. PLoS Comput Biol 2011 Oct;7(10):e1002199 [FREE Full text] [doi: 10.1371/journal.pcbi.1002199] [Medline: 22022249]
25. Aslam AA, Tsou M, Spitzberg BH, An L, Gawron JM, Gupta DK, et al. The reliability of tweets as a supplementary method of seasonal influenza surveillance. J Med Internet Res 2014;16(11):e250 [FREE Full text] [doi: 10.2196/jmir.3532] [Medline: 25406040]

26. Ji X, Chun SA, Wei Z, Geller J. Twitter sentiment classification for measuring public health concerns. Soc Netw Anal Min 2015 May 12;5(1):13. [doi: 10.1007/s13278-015-0253-5]
27. Weeg C, Schwartz HA, Hill S, Merchant RM, Arango C, Ungar L. Using twitter to measure public discussion of diseases: a case study. JMIR Public Health Surveill 2015;1(1):e6 [FREE Full text] [doi: 10.2196/publichealth.3953] [Medline: 26925459]
28. Alvaro N, Conway M, Doan S, Lofi C, Overington J, Collier N. Crowdsourcing twitter annotations to identify first-hand experiences of prescription drug use. J Biomed Inform 2015 Dec;58:280-287 [FREE Full text] [doi: 10.1016/j.jbi.2015.11.004] [Medline: 26556646]
29. Robertson C, Yee L. Avian influenza risk surveillance in North America with online media. PLoS One 2016;11(11):e0165688 [FREE Full text] [doi: 10.1371/journal.pone.0165688] [Medline: 27880777]
30. Taddio A, Ipp M, Thivakaran S, Jamal A, Parikh C, Smart S, et al. Survey of the prevalence of immunization non-compliance due to needle fears in children and adults. Vaccine 2012 Jul 06;30(32):4807-4812. [doi: 10.1016/j.vaccine.2012.05.011] [Medline: 22617633]
31. Wheelock A, Parand A, Rigole B, Thomson A, Miraldo M, Vincent C, et al. Socio-psychological factors driving adult vaccination: a qualitative study. PLoS One 2014;9(12):e113503 [FREE Full text] [doi: 10.1371/journal.pone.0113503] [Medline: 25490542]
32. Rossmann BE, Rench MA, Montesinos DP, Healy CM. Acceptability of immunization in adult contacts of infants: possibility of expanding platforms to increase adult vaccine uptake. Vaccine 2014 May 07;32(22):2540-2545. [doi: 10.1016/j.vaccine.2014.03.056] [Medline: 24681227]
33. Kim N, Mountain TP. Role of non-traditional locations for seasonal flu vaccination: empirical evidence and evaluation. Vaccine 2017 May 19;35(22):2943-2948. [doi: 10.1016/j.vaccine.2017.04.023] [Medline: 28433328]
34. Prematunge C, Corace K, McCarthy A, Nair RC, Roth V, Suh KN, et al. Qualitative motivators and barriers to pandemic vs. seasonal influenza vaccination among healthcare workers: a content analysis. Vaccine 2014 Dec 12;32(52):7128-7134. [doi: 10.1016/j.vaccine.2014.10.023] [Medline: 25454876]
35. Kelso JM. Influenza vaccine and egg allergy: nearing the end of an evidence-based journey. J Allergy Clin Immunol Pract 2015;3(1):140-141. [doi: 10.1016/j.jaip.2014.08.011] [Medline: 25577641]
36. Zerbo O, Qian Y, Yoshida C, Fireman BH, Klein NP, Croen LA. Association between influenza infection and vaccination during pregnancy and risk of autism spectrum disorder. JAMA Pediatr 2017 Jan 02;171(1):e163609. [doi: 10.1001/jamapediatrics.2016.3609] [Medline: 27893896]
37. Maher L, Dawson A, Wiley K, Hope K, Torvaldsen S, Lawrence G, et al. Influenza vaccination during pregnancy: a qualitative study of the knowledge, attitudes, beliefs, and practices of general practitioners in Central and South-Western Sydney. BMC Fam Pract 2014 May 23;15:102 [FREE Full text] [doi: 10.1186/1471-2296-15-102] [Medline: 24884996]
38. Chamberlain A, Seib K, Ault K, Orenstein W, Frew P, Malik F, et al. Factors associated with intention to receive influenza and tetanus, diphtheria, and acellular pertussis (Tdap) vaccines during pregnancy: a focus on vaccine hesitancy and perceptions of disease severity and vaccine safety. PLoS Curr 2015 Feb 25;7:- [FREE Full text] [doi: 10.1371/currents.outbreaks.d37b61bceebae5a7a06d40a301cfa819] [Medline: 25789203]
39. Bettinger JA, Greyson D, Money D. Attitudes and beliefs of pregnant women and new mothers regarding influenza vaccination in British Columbia. J Obstet Gynaecol Can 2016 Nov;38(11):1045-1052. [doi: 10.1016/j.jogc.2016.08.004] [Medline: 27969559]
40. Lau JT, Au DW, Tsui HY, Choi KC. Prevalence and determinants of influenza vaccination in the Hong Kong Chinese adult population. Am J Infect Control 2012 Sep;40(7):e225-e227. [doi: 10.1016/j.ajic.2012.01.036] [Medline: 22743046]
41. Lv M, Fang R, Wu J, Pang X, Deng Y, Lei T, et al. The free vaccination policy of influenza in Beijing, China: The vaccine coverage and its associated factors. Vaccine 2016 Apr 19;34(18):2135-2140. [doi: 10.1016/j.vaccine.2016.02.032] [Medline: 26917011]
42. Gargano L, Underwood N, Sales J, Seib K, Morfaw C, Murray D, et al. Influence of sources of information about influenza vaccine on parental attitudes and adolescent vaccine receipt. Hum Vaccin Immunother 2015;11(7):1641-1647 [FREE Full text] [doi: 10.1080/21645515.2015.1038445] [Medline: 25996686]
43. Admon L, Haefner JK, Kolenic GE, Chang T, Davis MM, Moniz MH. Recruiting pregnant patients for survey research: a head to head comparison of social media-based versus clinic-based approaches. J Med Internet Res 2016 Dec 21;18(12):e326 [FREE Full text] [doi: 10.2196/jmir.6593] [Medline: 28003174]
44. Yang M, Tan EC. Web-based and telephone surveys to assess public perception toward the national health insurance in taiwan: a comparison of cost and results. Interact J Med Res 2015 Apr 17;4(2):e9 [FREE Full text] [doi: 10.2196/ijmr.4090] [Medline: 25886989]
45. Santillana M, Nguyen AT, Dredze M, Paul MJ, Nsoesie EO, Brownstein JS. Combining search, social media, and traditional data sources to improve influenza surveillance. PLoS Comput Biol 2015 Oct;11(10):e1004513 [FREE Full text] [doi: 10.1371/journal.pcbi.1004513] [Medline: 26513245]
46. Guthmann J, Fonteneau L, Bonmarin I, Lévy-Bruhl D. Influenza vaccination coverage one year after the A(H1N1) influenza pandemic, France, 2010-2011. Vaccine 2012 Feb 01;30(6):995-997. [doi: 10.1016/j.vaccine.2011.12.011] [Medline: 22178520]

47. Allen C, Tsou M, Aslam A, Nagel A, Gawron J. Applying GIS and machine learning methods to Twitter data for multiscale surveillance of influenza. PLoS One 2016;11(7):e0157734 [FREE Full text] [doi: 10.1371/journal.pone.0157734] [Medline: 27455108]
48. Tuarob S, Tucker C, Salathe M, Ram N. Discovering health-related knowledge in social media using ensembles of heterogeneous features. In: Proceedings of the 22nd ACM International Conference on Information and Knowledge Management (CIKM 2013). New York: ACM; 2013 Presented at: CIKM 2013; October 27 - November 01, 2013; San Francisco p. 1685-1690. [doi: 10.1145/2505515.2505629]
49. Jain VK, Kumar S. An effective approach to track levels of influenza-A (H1N1) pandemic in India using twitter. Procedia Computer Science 2015;70:801-807. [doi: 10.1016/j.procs.2015.10.120]
50. Prieto VM, Matos S, Álvarez M, Cacheda F, Oliveira JL. Twitter: a good place to detect health conditions. PLoS One 2014;9(1):e86191 [FREE Full text] [doi: 10.1371/journal.pone.0086191] [Medline: 24489699]
51. Aronson A. Effective mapping of biomedical text to the UMLS Metathesaurus: the MetaMap program. Proc AMIA Symp 2001:17-21 [FREE Full text] [Medline: 11825149]
52. Cohen J. A coefficient of agreement for nominal scales. Educ Psychol Meas 2016 Jul 02;20(1):37-46. [doi: 10.1177/001316446002000104]
53. De Marnefe MC, Manning CD. Stanford University. Stanford typed dependencies manual URL: https://nlp.stanford.edu/software/dependencies_manual.pdf [accessed 2017-08-18] [WebCite Cache ID 6so2F3bX4]
54. Manning C, Surdeanu M, Bauer J, Finkel J, Bethard S, McClosky D. The Stanford coreNLP natural language processing toolkit. In: Proceedings of 52nd Annual Meeting of the Association for Computational Linguistics.: ACL; 2014 Presented at: System Demonstrations; June 23-24, 2014; Baltimore p. 55-60 URL: http://www.aclweb.org/anthology/P/P14/P14-5010 [doi: 10.3115/v1/P14-5010]
55. Chang C, Lin C. LIBSVM. ACM Trans Intell Syst Technol 2011 Apr 01;2(3):1-27. [doi: 10.1145/1961189.1961199]
56. Blei DM. Probabilistic topic models. Commun ACM 2012 Apr 01;55(4):77. [doi: 10.1145/2133806.2133826]
57. Blei D, Ng A, Jordan M. Latent dirichlet allocation. J Mach Learn Res 2003;3(Jan):993-1022.
58. Mallet. 2002. MALLET: a machine learning for language toolkit URL: http://mallet.cs.umass.edu/ [accessed 2016-11-16] [WebCite Cache ID 6m3SO8CIp]
59. Graham S, Weingart S, Milligan I. The Programming Historian. 2012. Getting started with topic modeling and MALLET URL: https://programminghistorian.org/lessons/topic-modeling-and-mallet [accessed 2017-08-18] [WebCite Cache ID 6so8n0jwd]
60. Han YK, Michie S, Potts HW, Rubin GJ. Predictors of influenza vaccine uptake during the 2009/10 influenza A H1N1v ('swine flu') pandemic: results from five national surveys in the United Kingdom. Prev Med 2016 Mar;84:57-61 [FREE Full text] [doi: 10.1016/j.ypmed.2015.12.018] [Medline: 26757401]
61. Bíró A. Determinants of H1N1 vaccination uptake in England. Prev Med 2013 Aug;57(2):140-142. [doi: 10.1016/j.ypmed.2013.04.017] [Medline: 23648526]
62. Johnson DR, Nichol KL, Lipczynski K. Barriers to adult immunization. Am J Med 2008 Jul;121(7 Suppl 2):S28-S35. [doi: 10.1016/j.amjmed.2008.05.005] [Medline: 18589065]
63. Giese C, Mereckiene J, Danis K, O'Donnell J, O'Flanagan D, Cotter S. Low vaccination coverage for seasonal influenza and pneumococcal disease among adults at-risk and health care workers in Ireland, 2013: The key role of GPs in recommending vaccination. Vaccine 2016 Jul 12;34(32):3657-3662. [doi: 10.1016/j.vaccine.2016.05.028] [Medline: 27255466]
64. Campbell H, Edwards A, Letley L, Bedford H, Ramsay M, Yarwood J. Changing attitudes to childhood immunisation in English parents. Vaccine 2017 May 19;35(22):2979-2985. [doi: 10.1016/j.vaccine.2017.03.089] [Medline: 28442229]
65. He L, Liao Q, Huang Y, Feng S, Zhuang X. Parents' perception and their decision on their children's vaccination against seasonal influenza in Guangzhou. Chin Med J (Engl) 2015 Feb 05;128(3):327-341. [doi: 10.4103/0366-6999.150099] [Medline: 25635428]
66. Chan T, Fu Y, Wang D, Chuang J. Determinants of receiving the pandemic (H1N1) 2009 vaccine and intention to receive the seasonal influenza vaccine in Taiwan. PLoS One 2014;9(6):e101083 [FREE Full text] [doi: 10.1371/journal.pone.0101083] [Medline: 24971941]
67. Taddio A, McMurtry CM, Shah V, Riddell RP, Chambers CT, Noel M, HELPinKids&Adults. Reducing pain during vaccine injections: clinical practice guideline. CMAJ 2015 Sep 22;187(13):975-982 [FREE Full text] [doi: 10.1503/cmaj.150391] [Medline: 26303247]
68. Taddio A, McMurtry CM, Shah V, Yoon EW, Uleryk E, Pillai RR, HELPinKids&Adults Team. Methodology for knowledge synthesis of the management of vaccination pain and needle fear. Clin J Pain 2015 Oct;31(10 Suppl):S12-S19 [FREE Full text] [doi: 10.1097/AJP.0000000000000263] [Medline: 26352917]
69. Eppes C, Wu A, You W, Cameron KA, Garcia P, Grobman W. Barriers to influenza vaccination among pregnant women. Vaccine 2013 Jun 12;31(27):2874-2878. [doi: 10.1016/j.vaccine.2013.04.031] [Medline: 23623863]
70. Evanovich LL, Scott TM. Facilitating PBIS implementation: an administrator's guide to presenting the logic and steps to faculty and staff. Beyond Behavior 2016 Apr;25(1):4-8. [doi: 10.1177/107429561602500102]

Abbreviations

API: application programming interface
ILI: influenza-like illness
LDA: latent Dirichlet allocation
ML: machine learning
SVM: support vector machine
TF: term frequency
TF-IDF: term frequency-inverse document frequency
UMLS: unified medical language system

Edited by I Weber; submitted 25.01.17; peer-reviewed by L Ungar, Z Yin, D Zhang; comments to author 18.02.17; revised version received 09.06.17; accepted 26.07.17; published 12.09.17.

Please cite as:
Kagashe I, Yan Z, Suheryani I
Enhancing Seasonal Influenza Surveillance: Topic Analysis of Widely Used Medicinal Drugs Using Twitter Data
J Med Internet Res 2017;19(9):e315
URL: http://www.jmir.org/2017/9/e315/
doi:10.2196/jmir.7393
PMID:28899847

Original Paper

The User Knows What to Call It: Incorporating Patient Voice Through User-Contributed Tags on a Participatory Platform About Health Management

Annie T Chen[1], MSIS, PhD; Rachel M Carriere[2], MLS; Samantha Jan Kaplan[3], MLIS

[1]Department of Biomedical Informatics and Medical Education, University of Washington School of Medicine, Seattle, WA, United States

[2]EBSCO Information Services, Ipswich, MA, United States

[3]School of Information and Library Science, University of North Carolina Chapel Hill, Chapel Hill, NC, United States

Corresponding Author:
Annie T Chen, MSIS, PhD
Department of Biomedical Informatics and Medical Education
University of Washington School of Medicine
850 Republican Street, Box 358047, C238
Seattle, WA, 98109
United States
Phone: 1 206 221 9218
Fax: 1 206 221 2671
Email: atchen@uw.edu

Abstract

Background: Body listening, described as the act of paying attention to the body's signals and cues, can be an important component of long-term health management.

Objective: The aim of this study was to introduce and evaluate the Body Listening Project, an innovative effort to engage the public in the creation of a public resource—to leverage collective wisdom in the health domain. This project involved a website where people could contribute their experiences of and dialogue with others concerning body listening and self-management. This article presents an analysis of the tags contributed, with a focus on the value of these tags for knowledge organization and incorporation into consumer-friendly health information retrieval systems.

Methods: First, we performed content analysis of the tags contributed, identifying a set of categories and refining the relational structure of the categories to develop a preliminary classification scheme, the Body Listening and Self-Management Taxonomy. Second, we compared the concepts in the Body Listening and Self-Management Taxonomy with concepts that were automatically identified from an extant health knowledge resource, the Unified Medical Language System (UMLS), to better characterize the information that participants contributed. Third, we employed visualization techniques to explore the concept space of the tags. A correlation matrix, based on the extent to which categories tended to be assigned to the same tags, was used to study the interrelatedness of the taxonomy categories. Then a network visualization was used to investigate structural relationships among the categories in the taxonomy.

Results: First, we proposed a taxonomy called the Body Listening and Self-Management Taxonomy, with four meta-level categories: (1) health management strategies, (2) concepts and states, (3) influencers, and (4) health-related information behavior. This taxonomy could inform future efforts to organize knowledge and content of this subject matter. Second, we compared the categories from this taxonomy with the UMLS concepts that were identified. Though the UMLS offers benefits such as speed and breadth of coverage, the Body Listening and Self-Management Taxonomy is more consumer-centric. Third, the correlation matrix and network visualization demonstrated that there are natural areas of ambiguity and semantic relatedness in the meanings of the concepts in the Body Listening and Self-Management Taxonomy. Use of these visualizations can be helpful in practice settings, to help library and information science practitioners understand and resolve potential challenges in classification; in research, to characterize the structure of the conceptual space of health management; and in the development of consumer-centric health information retrieval systems.

Conclusions: A participatory platform can be employed to collect data concerning patient experiences of health management, which can in turn be used to develop new health knowledge resources or augment existing ones, as well as be incorporated into consumer-centric health information systems.

(J Med Internet Res 2017;19(9):e292) doi:10.2196/jmir.7673

KEYWORDS

collaborative tagging; folksonomy; knowledge organization; self-management; body listening; body awareness

Introduction

Background

In recent years, there has been increased interest in Web-based platforms that aim to derive value from user participation through crowdsourcing, collaborative, and participatory frameworks. Efforts to leverage "collective intelligence" include the collective authoring of Wikipedia content, shared tagging of photos on Flickr, sharing of bookmarks on Del.icio.us, and collective annotation of museum artifacts [1-5]. Collaborative tagging, a practice in which users add meta-data to shared content, can be useful when personnel are not readily available to perform classification tasks [6], as a channel for nonprofessional catalogers to participate in meta-data creation [7], and as an approach to organize knowledge by users' own language [8].

Due to their collaborative and ad hoc nature, tagging systems inherently lack the essential properties characterizing controlled vocabularies, and low precision and lack of collocation are common issues [9]. However, this very nature also provides support for multiplicity of perspectives, collective interpretation, sense-making, and meaning production; and promotion of a collaborative, democratic, and participatory style of knowledge construction and organization [10-12].

Collaborative tagging presents a special opportunity in biomedical knowledge organization. There are numerous knowledge resources such as the Unified Medical Language System (UMLS) Metathesaurus [13] and the Systematized Nomenclature of Medicine-Clinical Terms (SNOMED-CT) [14], which facilitate inference in biomedical and clinical domains. However, there has been increased awareness that health consumers possess a type of expertise that is different from that of clinicians and that user-generated content can be a valuable source of health knowledge [15]. As such, there has been work that has employed social media data to improve existing knowledge resources, including the use of PatientsLikeMe data to augment the open access and collaborative Consumer Health Vocabulary [16] and the use of community-generated text to map professional medical terms to their consumer equivalents [17]. Other research has investigated the overlap between social media data sources such as PatientsLikeMe and YouTube and SNOMED-CT [18,19]. It has also been shown that search log data share similarities to folksonomy tags and can be used to improve controlled vocabularies and information retrieval [20].

The use of social media data to augment professional controlled vocabularies is important work. In this study, we concentrated our efforts in a different direction: to employ collective intelligence to develop a knowledge resource focused on patients' health management strategies. To set the context for the knowledge resource that we aimed to build, we now review extant literature on self-management of chronic illness, body listening, and related terminology.

Self-Management of Chronic Illness and the Importance of Body Listening

Over the course of a chronic illness, people learn to manage their health in different ways. Self-management of chronic illness has been characterized as a dynamic and daily experience involving three main categories of processes: focusing on illness needs, activating resources, and living with a chronic illness [21]. It has been argued that though there is considerable extant research on self-management barriers and facilitators, the developmental patterns and sustainability of self-management over time remain largely unknown [22]. In this study, we set out to investigate how people form and acquire self-management skills, particularly those relating to body listening and body awareness.

Body awareness, defined as the ability to recognize subtle body cues, can be helpful in the management of many conditions including chronic low back pain, congestive heart failure, chronic renal failure, and irritable bowel syndrome [23]. In the context of fibromyalgia, patients have reported learning over time what their pain triggers were, coming to understand what foods they were sensitive to, and when they had hit their limit and needed to rest [24].

Combining an awareness of embodied experience with information from test results, that is, combining one's own knowledge with a biomedical understanding of a condition, has also been referred to as "knowing one's body" [25]. Body listening has also been described as the "subprocesses of physical self-assessment and applying a personal filter through which to interpret that information" [26] (p. 265).

Body awareness and body listening also share similarities with concepts such as self-awareness and self-monitoring. For example, previous research has investigated the self-awareness of the cues, sensations, and circumstances that people with diabetes associate with hypoglycemia, euglycemia, and hyperglycemia, and the types of strategies that they used to tune in to these body cues [27]. Self-monitoring has been defined as having two components: (1) awareness of bodily symptoms, sensations, daily activities, and cognitive processes; and (2) measurements, recordings, and observations that inform cognition or provide information for independent action or consultation with care providers [28].

The importance of attending to body cues in health management suggests that a greater understanding of the ways in which people engage in these activities is warranted. Moreover, patients and clinicians may have different ways of describing their symptom experiences [29]. Thus, if the knowledge that patients acquire over time could be effectively captured, this information might be incorporated into health knowledge resources and shared on a wider scale. To take a step to fill this need, we developed a platform where people could post their experiences concerning body listening and how they learned or were learning to do it, in the hope that the data collected could later be used to develop a system for organizing and exploring knowledge relating to body listening.

Research Questions

This study was based on a premise of the value of coconstruction of knowledge. We developed a platform on which project team members and study participants could collectively engage in discussion on topics relating to body listening and, more broadly, health management. We targeted chronic conditions requiring self-care in which patients might share symptomatology, with a particular focus on chronic pain. These conditions included chronic pain, fibromyalgia, arthritis, and multiple sclerosis and common comorbidities such as chronic fatigue syndrome and irritable bowel syndrome.

Over the course of 10 weeks, participants engaged in a moderated discussion of topics relating to body listening. As they engaged in discussion, participants were also encouraged to provide tags that described the content that they shared. This article investigates whether asking participants to provide tags can add value to the data that they contribute and what the potential implications of this value might be to knowledge organization. We investigated three main research questions:

RQ1. What types of subject matter were represented in the tags?

RQ2. How does concept coverage in the Body Listening and Self-Management Taxonomy, proposed in RQ1, compare with that of the UMLS Metathesaurus?

RQ3. To what extent were Body Listening and Self-Management Taxonomy categories assigned to the same tag, and what does this pattern of category cooccurrences suggest about the relationship between categories?

Methods

Data Collection

Study Platform and Content Development

This study involved a 10-week “Guided Exploration” in a discussion forum style space built using the Wordpress platform (Figure 1). The discussion forum itself was called the “ThinkSpace,” to emphasize the value of participants’ contributions. Data collection took place over the course of a 10-week period involving moderator-facilitated discussions of topics relating to body listening. Each week, a new theme was introduced, and each day, a new question relating to the theme was posted (Table 1).

Table 1. Guided Exploration schedule.

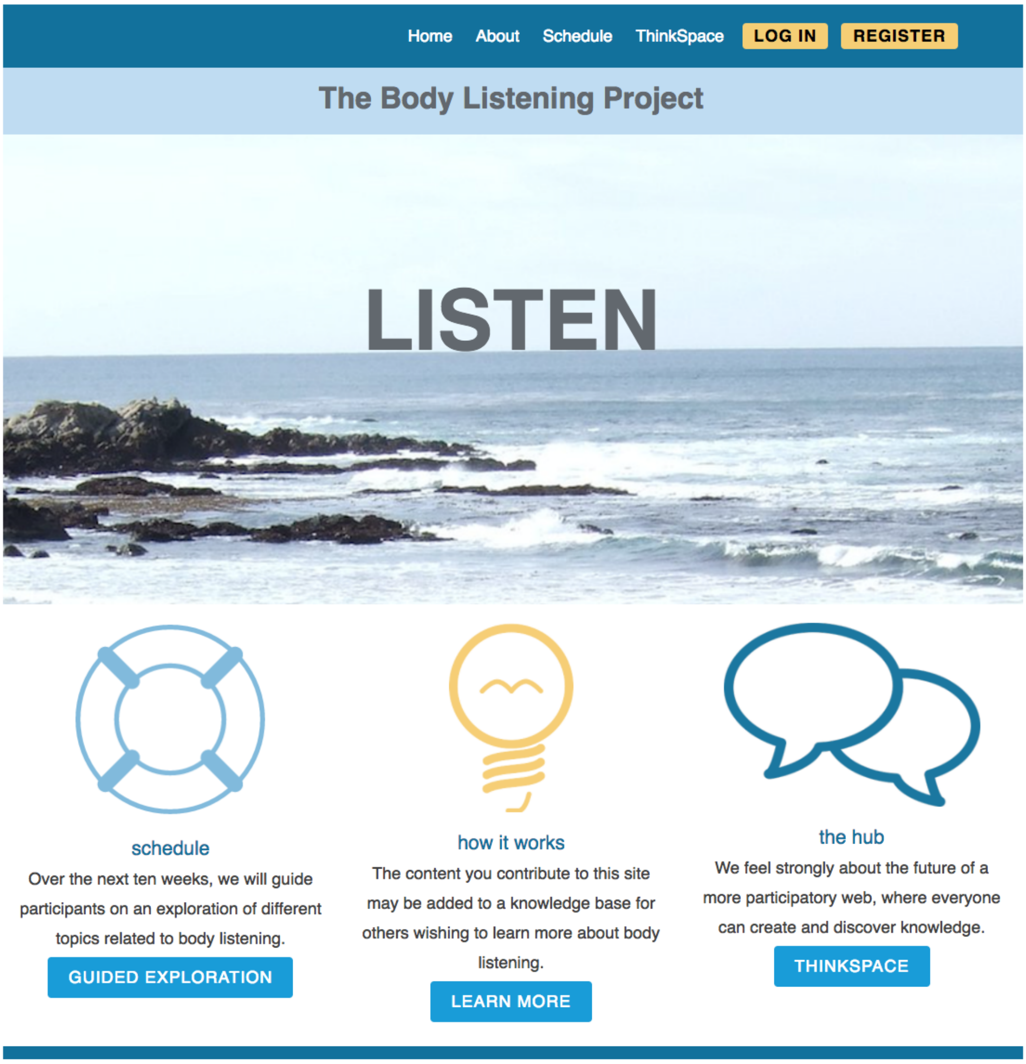
Week no.	Topic
1	Getting in touch with your body rhythms
2	Movement, energy, and fatigue
3	Food and environment
4	Pain management
5	Mood management
6	Sense-making and conveying what your body tells you in health care contexts
7	Conveying what your body tells you in life contexts
8	Tuning in to your body with arts-based techniques
9	Mindfulness as a way to get in touch with your body
10	The body as a vehicle for self-growth

Though there are processes that may be shared between individuals who engage in body listening and self-management, the set of experiences and perspectives is admittedly diverse. It was necessary to limit the topics covered so that they could be explored in the 10-week study period. The themes employed an experiential line of inquiry, and the content was tailored toward chronic conditions requiring self-care that overlapped in symptomatology, with chronic pain as the primary characteristic.

Selection of topics was based on a two-step process. At the outset, the first author developed a list of seed topics based on three sets of research literature: (1) self-management of chronic illness; (2) body listening and body awareness; and (3) fibromyalgia, arthritis, multiple sclerosis, and other chronic pain and rheumatological conditions. The moderators subsequently discussed these topics, added their own, and then finalized the set of topics to explore in the Guided Exploration. The content was collectively developed by the team through a series of biweekly meetings and the collaborative authoring of a shared document.

Moderators monitored the site regularly to engage the group and to respond to posts by site participants. Participants were encouraged to author posts and then add one or more tags after the main body of content. Participants were informed that the experiences they shared in the ThinkSpace would be considered data, and they gave consent as part of the account creation process. The study procedures were approved by the Institutional Review Board of the University of Washington School of Medicine.

Figure 1. The Body Listening Project index page.



Over the next ten weeks, we will guide participants on an exploration of different topics related to body listening.

The content you contribute to this site may be added to a knowledge base for others wishing to learn more about body listening.

We feel strongly about the future of a more participatory web, where everyone can create and discover knowledge.

The team comprised 10 people: a project manager, faculty advisors, moderators, a site administrator, and communications and outreach personnel. Members came from multiple disciplines including library and information science, biomedical informatics, nursing, and public health. Many team members also had experience with chronic illnesses.

Sample and Study Recruitment

The overall goal of this study was to increase our knowledge of body listening and body awareness, which can potentially be helpful in the management of many chronic conditions [23]. The study was open to anyone over the age of 21 years who was interested in, or wanted to contribute to, knowledge on the phenomenon of “body listening.” We employed two types of recruitment strategies. First, we employed recruitment strategies that were not focused on a given disease or condition, such as university and hospital listservs and posted flyers in high traffic areas such as university campuses and coffee shops.

Second, we recruited participants through multiple social media channels including Facebook, Reddit, Twitter, and health discussion forums. Our social media recruitment focused on particular groups of interest including groups focused on chronic pain, fibromyalgia, and discussion groups for commonly occurring comorbidities such as chronic fatigue syndrome and irritable bowel syndrome. We also reached out to people of different demographic characteristics including ethnicity and education level.

Data Analysis

RQ1. What Types of Subject Matter Were Represented in the Tags?

To analyze the tags that were collected, the posts were exported from the MySQL database where they were stored, and a PHP script was used to extract all of the tags. A hashtag was defined as a word or set of words strung together without spaces, preceded by a hash or pound sign (#). Hashtags can be useful for assisting users to find content of interest. There were some malformed tags—cases in which multi-word tags included spaces, such as “#cell phone alarm,” and instances in which tags started and ended with pound signs, such as “#meditation/prayer time#.” Twelve such errors were caught through manual lookup, and the tags’ formats were corrected for inclusion in the analysis. We did not correct for spelling errors and spelling variants, as we believed that this would be useful information to preserve.

Though it was expected that many participants might have familiarity with formatting guidelines for social media and microblog hashtagging, we developed an informational page that explained to participants how to author content. A link to this “Contribution How-To” page was available from the main ThinkSpace page. Sample tags were provided to demonstrate tag content and formatting. However, to avoid limiting the participants’ conceptualization of what a tag was, an explicit definition of a tag was not included as part of this page.

To understand the nature of the tags that were used, the tags were manually assigned to one or more categories by the first two authors, and a relational hierarchy of the codes developed, as follows. First, the two coders independently assigned categories to a subset of the data (n=100). They discussed these codes and categories until they reached agreement about the categories and the hierarchical relationship between them. Then they coded a new subset (n=100) and calculated three different types of inter-rater reliability (IRR). A more extensive discussion of the types of IRR used appears in a prior work [30], but a summary of the types of IRR appears in Multimedia Appendix 1. In this round of coding, they achieved a high level of agreement, so the second author proceeded to code the rest of the tags. The third author also took part in the decision-making process concerning the relationships between the categories.

This taxonomy was later slightly refined to improve the clarity of category definitions, as well as to increase the inclusiveness and logical consistency of the framework by renaming the meta-category, “gathering and conveying information,” to “health-related information behavior,” and adding the “health-related information behavior” category. All three authors took part in the discussion to refine the taxonomy. The taxonomy is presented in the Results section.

There were situations in which the meanings of the tags were unclear. For example, #isolation may refer to a state, a treatment modality, or exercise. Ambiguous tags were analyzed and categorized within the context of the posts in which they occurred. A list of the tags, along with their classifications, appears in Multimedia Appendix 2. If a tag was not formatted correctly in the original contribution, the original malformed tag appears in parentheses.

RQ2. How Does Concept Coverage in the Body Listening and Self-Management Taxonomy Compare With That of the Unified Medical Language System (UMLS) Metathesaurus?

To investigate the extent to which tags expressed concepts covered in the UMLS Metathesaurus, we manually separated all multi-word tags into their individual word components and used the MetaMap API to identify relevant UMLS concepts in these tags [31]. We then performed error analysis of a subset of the tags for which MetaMap identified relevant concepts (n=200) to evaluate the degree to which the identified concepts captured the meaning of the specified tags. Identified concepts were considered “correct” if they captured the essence of the meaning of the corresponding tags, “incorrect” if the meaning of the identified concepts did not match the meaning of the tag, and “incomplete” if there were critical parts of the meaning that were not captured. We also identified the sources of error if an identified concept was labeled “incorrect” or “incomplete.”

To visually compare the concept coverage, we generated category frequency distributions for both the Body Listening and Self-Management Taxonomy and for the UMLS using Python.

RQ3. To What Extent Were Body Listening and Self-Management Taxonomy Categories Assigned to the Same Tag, and What Does This Pattern of Category Cooccurrences Suggest About the Relationship Between Categories?

In this research question, we employ visualizations to explore the concept space of the contributed tags. A correlation matrix, based on the extent to which categories tended to be assigned to the same tags, was used to study the interrelatedness of the taxonomy categories. Then, a network visualization was used to explore the concept space of the tags contributed and investigate the structural relationships among the categories in the taxonomy.

To examine conceptual overlap, we employed the multiple tag assignments to visualize the correlations between categories based on their tendencies to be assigned to the same tags. First, the tags and their category assignments were used to construct an N × N cooccurrence matrix based on the frequency at which categories were assigned the same tag and categories were considered to cooccur with themselves. This matrix was then converted and visualized as a diagonal correlation matrix using Python.

We also investigated how strongly categories were related to one another, both within and across meta-categories, as well as their overall prevalence within the tag corpus using a network visualization. Network analysis is used in numerous disciplines and has been used to analyze many naturally occurring phenomena including communication patterns in an emergency department [32], the spread of disease outbreaks [33], and the structure of research in a given subject area or discipline [34]. Other uses of network analysis include the analysis of structural relationships among entities [35] and identification of community structure [36]. Network structures can also be used to examine the organization of human semantic knowledge [37]. In this study, we employ it to examine the relationships between the categories in the Body Listening and Self-Management Taxonomy.

The open graph visualization software Gephi was used to visualize the network of tags [38]. We rendered a network in which each node represented a category in the taxonomy. The size of the node was based on the number of times that the category was assigned to a tag. The nodes were assigned colors based on their meta-categories, such that all nodes of the same meta-category shared the same color. Two nodes were connected if they were ever assigned to the same tag within the corpus, and the weight of the edges was determined by number of times they were assigned to the same tag. The ForceAtlas2 layout was used to visualize the network [39] and the Label Adjust algorithm to eliminate visual overlap in labels.

Results

Data Collection

Over the course of the 10-week study, 234 participants registered to participate. As expected, the range of conditions that participants reported was diverse, with greater emphasis on pain conditions, mental health conditions, food sensitivities, irritable bowel syndrome, chronic fatigue syndrome, and thyroid disorders—all areas which were emphasized in the social media recruitment. A total of 28 participants posted in the discussion forum. This participation pattern is consistent with prior research on discussion forums, in which those who participate are a fraction of the countless others that may be “lurking” [40].

Altogether, participants and moderators authored 431 posts and used 818 tags. Of these, 680 tags were unique, and the tags used more than once were used 2-6 times. Some tags such as emotional freedom technique (EFT) and traditional Chinese medicine (TCM) appeared both in their traditional and acronym forms. Approximately a quarter of the posts did not contain tags (n=114), and many posts that included tags featured multiple tags (n=197).

Though many tags were single words (n=163), others were combinations of words (n=517). Many expressed a concept, including #attitudeiseverything, #taketimetounderstand, and #nothavingmypaininadvance. This last tag expressed the concept of trying to be positive and dealing with situations as they occur, rather than worrying about pain that one might experience later. There were instances in which tags were assigned to multiple categories such as #nightpain, which was classified as symptoms and illnesses (IS), rhythms and schedules (RS), and influencers (IF).

Data Analysis

RQ1. What Types of Subject Matter Were Represented in the Tags?

The contributed tags were grouped into thematic categories and an initial relational structure for the categories developed, validated, and reported in prior work [30]. The validation procedure is summarized in Multimedia Appendix 1. In this research question, we focus on describing the proposed taxonomy and exploring its implications for the classification of subject matter concerning health management. There were four meta-level categories: health management strategies, concepts and states, influencers, and health-related information behavior (Table 2).

Health Management Strategies

This meta-category encompasses the wide variety of strategies that people may employ to manage their health. In addition to treatments (TM), people may utilize other strategies including exercise (E), diet and food (DF), coping or coping strategies (CS), and supplies and equipment (SE). Supplies and equipment was classified under health management strategies because participants would share examples of tools that they used to manage their health, such as #netipotforallergies. Strategies that did not clearly fall into any of the other categories were assigned the category health management strategy (HMS). Examples of this include #cellphonealarm and #taketimetorecover.

Concepts and States

Concepts and states comprised three categories: general concepts (CN), positive actions (PA), and mental states (MS). Certain concepts had an inherently positive orientation, and these were classified as positive actions. These included terms such as #advocatingforyourself and #transformation. There were others that referred to mental states, such as #acceptance, #compassion, and #nofear. Finally, concepts that were more generic, such as #journey, #balance, and #energy were simply classified as general concepts.

Influencers

Though one may not always be aware of factors that affect their body condition, nevertheless there are factors that influence the state of one’s body at any given point in time. The meta-category, influencers, was used to describe these factors and included two categories: outside factors influencing the body and mind (IF) and body rhythms and schedules (RS). Examples of the former included seasonal influences, such as #fourseasons and #pollenispainful, and examples of the latter included internal factors, such as #nightowl and #mostproductivetimes.

Table 2. Body Listening and Self-Management Taxonomy.

Meta-category			Description	Examples
	Code	Category		
Health management strategies				
	HMS	Health management strategy	Health management strategy	Cellphonealarm, taketimetorecover, fulltimebodymanager
	TM	Treatment or treatment strategy	Treatments (physical, psychological, other) for managing chronic illness	acupuncture, alexandertechnique, biofeedback, funkplunkmyselfoutside
	E	Exercise	Exercises and movements, as part of a treatment or not	Adductorstretch, walking, yamunabodyrolling, exercisewithchildren
	DF	Diet and food	Food, herb, and supplement consumption practices and principles	atkinsdiet, eliminationdiet(s)
	CS	Coping or coping strategies	Strategies for coping with pain or illness	prayer, affirmations, reframing
	SE	Supplies or equipment	Tools or equipment for tracking, treating, or coping with chronic pain or illness	Netipotforallergies, bodypillow
Concepts				
	CN	General concepts	General concepts relating to body listening	Attitudeiseverything, journey, balance, energy
	PA	Positive actions	Actions with a positive outlook or orientation	advocatingforyourself, transformation, pushingthrough
	MS	Mental states	Mindsets and approaches to manage life with pain or illness	acceptance, compassion, nofear
Influencers				
	IF	Influencers	Outside factors influencing the body or mind	Sickweatherchanges, fourseasons, hotweather, pollenispainful
	RS	Rhythms and schedules	Rhythms around and of the body	Nightowl, mostproductivetimes
Health-related information behavior				
	HIB	Health-related information behaviors	Health-related information behavior such as seeking and/or sharing information	findinghealthinformation
	HR	Health-related resources	Resources for gathering health-related information	Patientexperts, trustedsites
	SM	Self-monitoring	Becoming aware of body rhythms, treatment, or symptoms	Journaling, symptomtracker, sleepdigestionconnection, bodyclues
Other categories				
	IS	Symptoms, illnesses	Symptoms, illnesses, or diseases	Shakiness, sleepdisorders
	HC	Health care-related terms	Health care-related terms	doctors, massagetherapists, physicaltherapist
	CM	Communication and relationships	Interactions, communications, and relationships with others and the self	Energyfromothers, supportivebossesrock
	AD	Moderator or administrative content	Moderator instructions	Useanycoloryoulike, usebothhands

Health-Related Information Behavior

This meta-category comprised three categories, health-related information behaviors (HIB), health-related resources (HR), and self-monitoring (SM). Examples of health-related sources of information that participants mentioned include #patientexperts and #trustedsites. The category of self-monitoring included both monitoring that included devices, as well as self-monitoring and exercising awareness without devices, such as #bodyclues.

Health-related information behaviors was used to classify all other terms that fell under this meta-category but not specifically under any of the other two terms. In this study, these terms were extremely rare, but we believe that health-related information behaviors is an important category and that it could be further differentiated into additional categories. Thus, we have chosen to retain it as part of the framework.

It may be useful to consider the relationship between the two meta-categories, health-related information behavior and health management strategies. Health-related information can be used to manage one’s health, and in the context of this study, often was. However, having information as a separate meta-category was intended to emphasize that in these activities information played the central role and health was the supporting context.

For example, the tag #findinghealthinformation is primarily about information and secondarily about health. In situations in which both meanings were equally salient, a tag might be categorized under each meta-category. One example was #journaling, which was categorized as both self-monitoring (SM) and treatment or treatment strategy (TM).

Other

An additional set of categories did not fit under any meta-level category. These include symptoms and illnesses (IS), health care-related terms (HC), communication and relationships (CM), and moderator or administrative content (AD). Symptoms and illnesses help us to understand how participants view their own health, and health care-related terms and communication and relationships help us to understand elements of a person’s context that play important roles in their long-term health management. The last category, moderator or administrative content, was most prominent in Week 8: Tuning in to Your Body with Arts-Based Techniques, in which the moderator provided instructions for participants to explore and express their physiological sensations using arts-based techniques.

RQ2. How Does Concept Coverage in the Body Listening and Self-Management Taxonomy Compare With That of the UMLS Metathesaurus?

In this section, we compare the prevalence of category assignments based on the Body Listening and Self-Management Taxonomy (Figure 2) with the top 20 UMLS categories that were automatically identified using the Metamap API (Figure 3). The category prevalence includes multiple category assignments, meaning that for tags that were assigned to multiple categories in the Body Listening and Self-Management Taxonomy or in the UMLS, each category assignment is depicted in the frequency distributions.

Examining the prevalence of categories across the tag corpus can help us to identify subject matter areas in which the study was successful in eliciting information. In so doing, we may also identify important factors for people to consider as they address their health management needs. In the Body Listening and Self-Management Taxonomy, health management strategies and more specifically, treatments, are one area in which participants contributed a great deal of content. A variety of treatments were mentioned, including widely known treatments such as #acupuncture and #biofeedback, lesser known techniques such as #alexandertechnique, and more original references such as #funkplunkmyselfoutside. Whether health management strategies are generally well-known or not, it can be useful to have them aggregated in one place. The more original ones may “speak to” or resonate with participants in a way that traditional resources may not.

Figure 2. Concept coverage in the Body Listening and Self-Management Taxonomy.

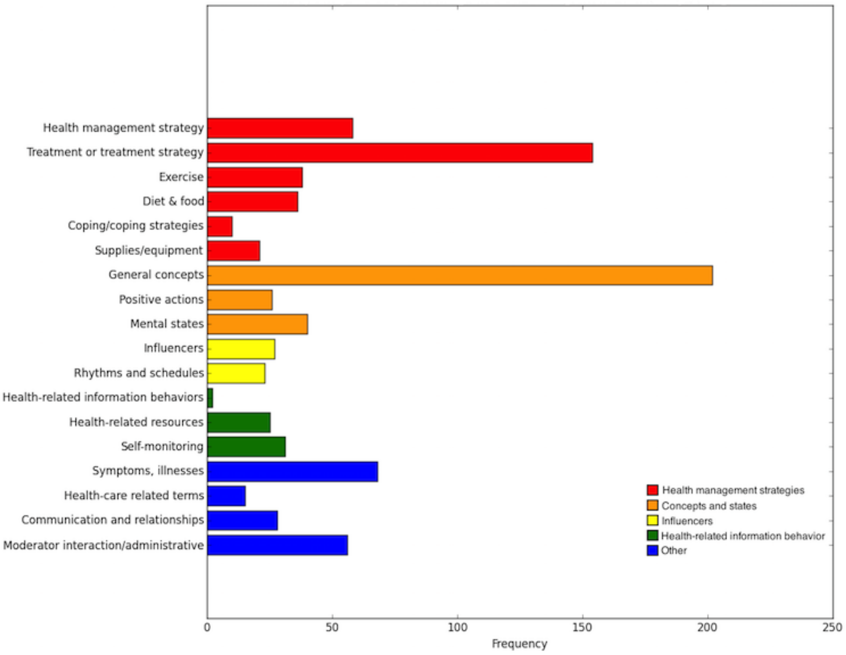
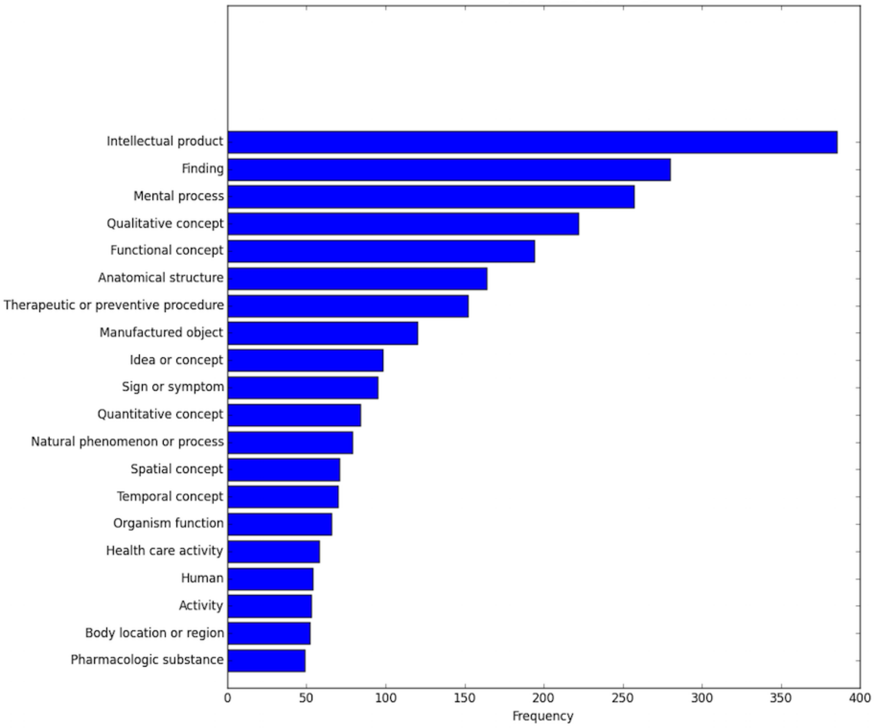


Figure 3. Concept coverage in the UMLS (Unified Medical Language System) Metathesaurus.



Another area in which there was extensive coverage was the category of general concepts. Though this category sounds fairly broad, we believe that additional study of the emergent concepts can lead to increased knowledge of important areas of self-management. The UMLS category of intellectual product perhaps roughly corresponds to the category of general concepts in the Body Listening and Self-Management Taxonomy.

We also investigated the extent to which concepts reflected in the tags were covered in the UMLS Metathesaurus. Altogether, 101 concepts were extracted; we show the highest frequency concepts in Figure 2. Some of the identified concepts are fairly clear in terms of their meaning, for example, sign or symptom and pharmacologic substance. Many, though, are too general (eg, “functional concept” or “qualitative concept”) to be useful to a health consumer searching for information.

The MetaMap API did not identify concepts for 11% (74/680) of the terms. We performed error analysis for a subset of the tags for which MetaMap identified UMLS concepts (n=200). Among these, MetaMap correctly identified one or more high-level classifications of the tags 60% (120/200) of the time, incorrectly identified the meaning 15.5% (31/200) of the time, and provided an incomplete set of relevant concepts 24.5% (49/200) of the time. Additionally, there were situations in which MetaMap correctly identified relevant concepts but produced noise through the identification of extraneous concepts. The primary sources of error (for incomplete or incorrect meaning) are shown in Table 3.

Table 3. Primary sources of error in automatic tag classification via the MetaMap API.

Issue	Example tag	Prevalence (%)
Not a strictly health-related concept	#supportivebossesrock	15
Meaning of an identified word is incorrect	#spiders (refers to the sensation on the skin, not the eukaryote)	12
Missing a concept	#noscentedcandles (missed “no” and “candle”)	11
Missing an interrogative word	#whatthebodywantstoeat	6
Missing a verb	#dowhatworksforyou	4
Is a health-related concept but is not included	#structuralintegration	4
Missing mental process	#distancethepain	3

Overall, automatic identification worked better for tags comprising one or two words. The more words that were involved, the more the tags tended to express a concept that was not well represented in its entirety by the identified concepts. Based on the error analysis in Table 3, areas for improvement of the UMLS include the integration of additional terms, including emergent terms such as structural integration; alternative and colloquial senses of words such as spiders on the skin; and contextual aspects of health management, such as workplace wellness.

These results suggest that though extant health knowledge resources might be used to characterize consumer-generated hashtags, additional work is necessary to support health consumers' information-seeking. Though this is a recognized problem that has received attention [16,41], there is still much that we can do to improve the UMLS coverage in terms of colloquial and patient or consumer-oriented language. Moreover, the results of this study suggest that, besides the addition of consumer-centric terms, there is also a need to add concepts to the UMLS. Lastly, the Body Listening and Self-Management Taxonomy could potentially fill an important gap in terms of providing conceptual categories that reflect patients' self-management strategies.

RQ3. To What Extent Were Body Listening and Self-Management Taxonomy Categories Assigned to the Same Tag, and What Does This Pattern of Category Cooccurrences Suggest About the Relationship Between Categories?

There were instances in which contributed tags were assigned to multiple categories. A total of 513 tags were assigned to only one category, and 167 tags received multiple assignments. In this section, we investigate the extent to which categories tended to be assigned to the same tag. The instances of multiple category assignment can help us to understand the naturally occurring relationships between taxonomy categories, the extent of their semantic overlap, and most importantly for practice, the extent of ambiguity that may be present during a manual process of classification.

We generated a diagonal correlation matrix to examine the tendency for categories to be assigned to the same tag (Figure 4). In this matrix, category combinations with a higher correlation, that is, a greater tendency to cooccur, are denoted in red. Health-related information behavior did not cooccur with any other categories during category assignment and thus, does not appear in the matrix.

Multi-category assignments were common among the health management strategies categories and among the three concepts and states categories. Examples of the former type of multi-category assignment include #hellerwork and #bodyrolling, both assigned to the exercise as well as treatment or treatment strategies category. Examples of overlap in the concepts and states categories include #compassion (CN, PA, and MS) and #letitgo (CA and PA). The pattern of correlations also reflects our own experiences in category assignment, in which we occasionally found it difficult to decide between treatment categories and concept categories because of the multiple senses that might be conveyed in tags. The cooccurrence of influencers with rhythms and schedules was also common. Examples of these include #fourseasons and #hotweather.

Another interesting aspect of these cooccurrences is the moderate correlations between the concepts and states categories and various categories within the health anagement strategies meta-category. The tags, #advocate4yourself (PA, HMS, and CN), which referred to advocating for oneself in workplace environments, and #allowingsadness (CN, PA, and TM), illustrate the importance of concepts and positive attitude in health management.

We also investigated the structural relationships between the categories—how strongly they were related to one another, both within and across meta-categories, as well as their overall prevalence within the tag corpus. We rendered a graph visualization as described in the Method section (Figure 5) and used the resulting visualization to engage in an interrogative dialogue concerning the relationships between the categories of the Body Listening and Self-Management Taxonomy. The larger the nodes, the greater their prevalence in the tag corpus, and the closer the category nodes were to each other in the network, the more times that they were assigned to the same tags. Nodes of the same color belonged to the same meta-category. As before, the health-related information behavior category does not appear because it did not cooccur with any other categories during category assignment.

The resulting visualization illuminates some interesting relationships between the categories. For example, most categories appear closer to other categories within the same meta-category. Categories under health-related information behavior are a notable exception. Though these three categories share conceptual similarity in that they all relate to how people interact with information, in the case of self-monitoring, there appeared to be a stronger conceptual connection with devices (supplies/equipment). Among the categories that were not subsumed under a meta-category, the position of symptoms and illnesses is particularly interesting. The proximity of the symptoms and illnesses node to rhythms and schedules underlines the importance of rhythms and temporal associations to symptoms. Examples of tags which were assigned to both categories include #reverseSAD, #daysnightsreversed, and #migraineuponwaking.

The proximity of categories to one another can provide insight into factors that may be important to consider together in body listening and health management. For example, this network shows that communication with others, mental states and health management strategies can be closely related. Whereas this may seem obvious, it provides further evidence that, in working with patients, it can be helpful for health care practitioners to consider contextual factors such as the patients' support network and mental state in assisting them to select, adopt and maintain health management strategies. The interrelatedness of the categories, as depicted in the figure, highlight the importance of not treating any aspect of a patient's care as if it occurs in a vacuum.

Figure 4. Co-occurrence of categories in tag classification.

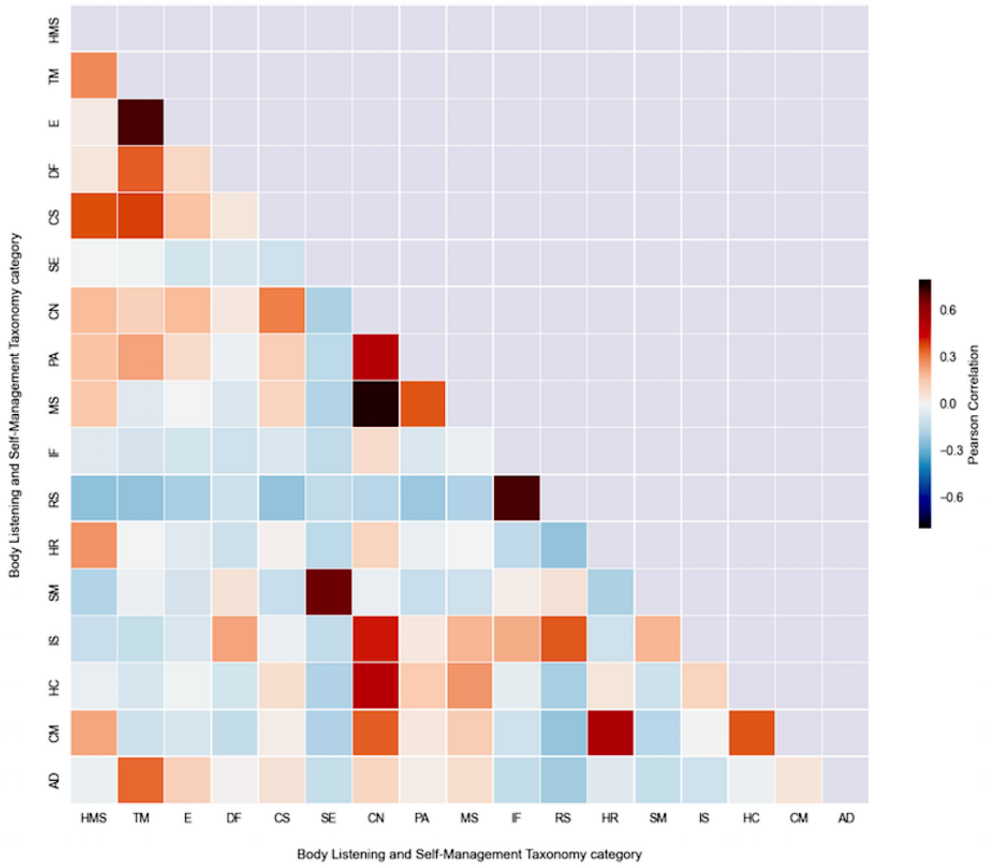
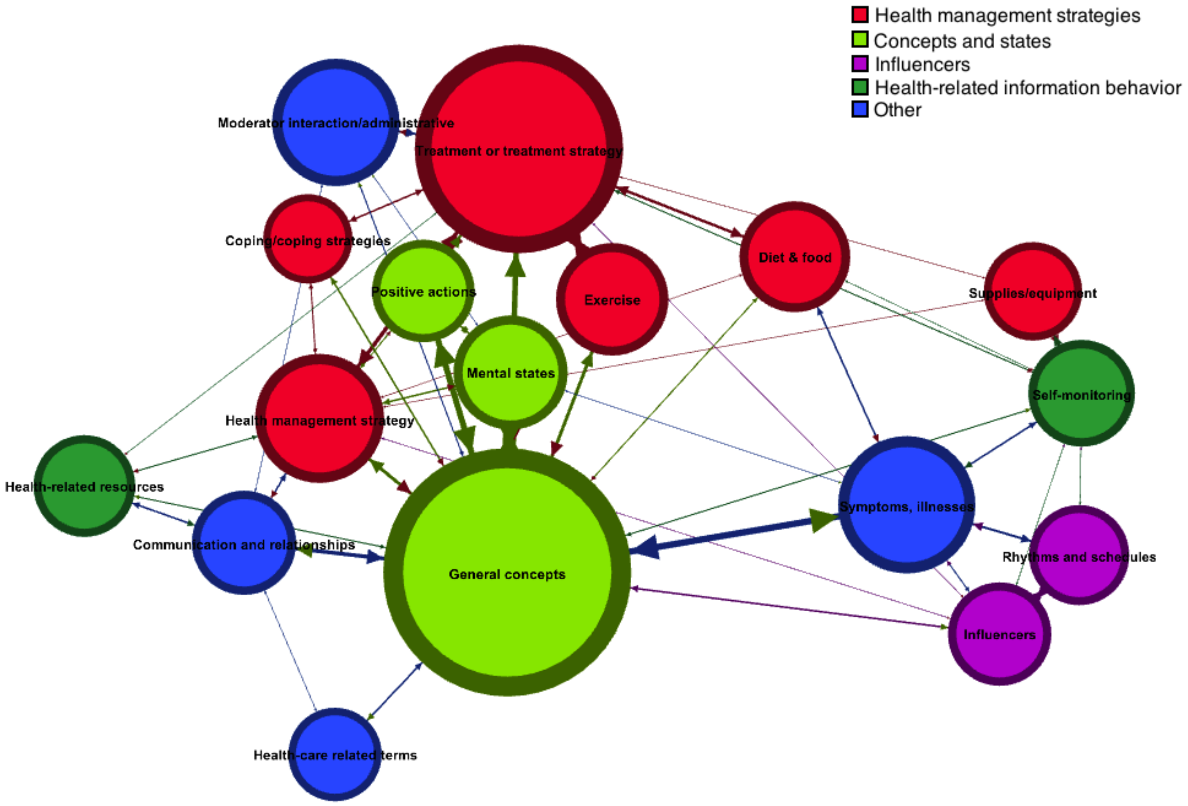


Figure 5. Body Listening and Self-Management Taxonomy as a network.



Discussion

Principal Contributions

This paper describes the development of a participatory platform for people to contribute their experiences of body listening and presents an analysis of the tags collected. This analysis makes several original contributions. First, we proposed a preliminary classification scheme, the Body Listening and Self-Management Taxonomy, for concepts associated with body listening. Second, we compared the concepts in the Body Listening and Self-Management Taxonomy with the concepts from the UMLS Metathesaurus that were automatically identified from the tags. This comparison served to characterize the information that patients contributed through the participatory platform and demonstrated that the tags offer important information concerning health management, particularly with regard to self-management strategies and contextual factors affecting health, such as seasonal and body rhythms, mental state, communication, and interpersonal relationships. Additionally, the comparison showed that the taxonomy could be used to classify patients' health management strategies more descriptively than could be achieved using UMLS concepts. Third, we examined the extent to which tags were assigned to multiple categories of the Body Listening and Self-Management Taxonomy, to gain insight into how the categories might be related to one another. Though the categories in any given meta-category generally appeared together, there were key areas of connectivity between the categories and meta-categories. Identifying these connections can help us to form a greater understanding of the need to consider factors affecting our health in context.

Considering the Value of the Data Collected Through the ThinkSpace

This study was intended as a starting point for building a public resource on body listening, and more broadly, self-management of chronic conditions. We used an approach combining user contribution of tags with subsequent curation to develop a taxonomy for use in the development of future knowledge resources.

This taxonomy is a preliminary step to improve the classification and retrieval of content that is presented to health consumers. We employ a two-step process that offers health consumers the opportunity for greater involvement by enabling them to not only provide data but also influence how it is organized through the tags that they provide. However, folksonomies come with certain issues, such as a preponderance of terms with different levels of specificity. Thus, we added an additional layer of human curation and classification to create a schema that is emergent from the data, and thus, more consumer-centric.

One important implication from the study results concerns the subject matter content of the contributed tags and the manner in which this data might be utilized. The ThinkSpace resulted in tags that focused on health management strategies, concepts, and contextual information about people's health experiences. These tags provide insight into the actions that patients take to manage their health, the realities that they face as they address their health issues, and the resources that they have access to, from their perspectives. Thus, the tags might be incorporated into a system that patients, caregivers, and health consumers in general can use to find information about their health needs and interests.

Though systems that provide health information to consumers do exist, they are generally focused around traditional medical concepts. For example, the Medline Plus interface currently includes section headings and subheadings such as "Health topics: find information on health, wellness, disorders, and conditions," "Drugs and supplements: learn about prescription drugs, over-the-counter medicines, herbs, and supplements," and "Medical encyclopedia: articles and images for diseases, symptoms, tests, treatments." Incorporating tags used in the ThinkSpace, such as #funkplunkmyselfoutside, #exercisewithchildren, and #supportivebossesrock, or the concepts expressed therein, may resonate with health information seekers as well as serve as an alternative channel through which to access information. In addition, our findings suggest that meta-data such as body rhythms, external influences, concepts, and attitudes could potentially serve as entry points for finding information. Concepts in the rhythms and schedules category, such as having pain at night and difficulties due to hot weather may resonate with patients but are less likely to appear as indices or headings in traditional consumer health information systems. Yet they are nevertheless important to health management. Thus, user-contributed tags could be incorporated into health information retrieval systems to make them more user-friendly and intuitive.

Facilitating Tag Classification and Taxonomy Enrichment

In this study, we classified the tags using two methods: manual assignment and automated identification. Manual assignment led to a set of categories that was more specific to body listening and self-management than those automatically identified through the MetaMap API, but manual classification can be time-consuming. Thus, we now consider ways to improve the classification methods employed in this study.

First, it might be useful to try to reduce the need for additional labor by engaging contributors in the creation of a taxonomy at the time of tag creation and use, which may also improve the specificity of tags and participants' memories of the content shared [42]. Examples of hybrid taxonomy-folksonomy approaches exist in the literature [43,44]. We might also assist users to tag content by providing recommendations generated through a variety of techniques, including tag cooccurrence, content-based, graph-based, and clustering- or topic-based methods [45]. However, it is important to consider whether the provision of tags might stifle creativity and prevent users from making a greater effort to fully elucidate their thoughts using the most appropriate tag. Finally, extant resources might be employed to automatically categorize tags. In previous literature, resources such as YAGO and WordNet have been used to categorize tags [46], to determine semantic relatedness [47], and to turn a folksonomy into a concept hierarchy [48].

Limitations

This study has various limitations. First, the duration of this study was 10 weeks, and thus topic coverage was limited to the amount of content that could be covered within that span of time. Given this consideration, the seed topics and the social media recruitment strategy were tailored to focus on a set of conditions with a shared symptomatology. Thus, the health management strategies that are reflected in the data are likely to primarily reflect the interests of this population, and it is necessary to conduct other studies with other target populations to better understand their health management needs.

In addition, the size of the tag corpus was small compared with larger and longer-term social networks such as Flickr and Del.icio.us, and the tag base had not reached a point of stabilization. In the future, it would be useful to employ additional strategies to increase the body of tags, which would most likely also result in revisions to the proposed taxonomy. These techniques might include using the contributed tags as seed terms to identify other relevant terms and categories from conventional knowledge resources or conducting additional Guided Explorations focused on related topics and/or targeting different populations.

Finally, this study informed participants from the outset that all data posted in the ThinkSpace would be public. Not everyone is willing to post in public spaces, and thus, the sample likely reflects this bias. However, we believe that the data contributed serves a valuable function to the public because even if people do not post, many of those who registered and also those who did not likely engaged with the content in their own way and were influenced by it. In addition, previous research has shown that there are different types of users and that a minority percentage of "verbose" taggers can produce results that match and even outperform the semantics from an entire dataset [49].

Acknowledgments

The authors would like to thank the study participants and pilot participants for their willingness to contribute and all of the members of the Body Listening Project research team for their assistance in the execution of this study.

Authors' Contributions

All three authors, along with the rest of the Body Listening Project research team, collaboratively developed the platform and collected the data. AC and RC performed the bulk of the data analysis for Research Question 1 and developed the initial version of the taxonomy, which was refined collaboratively by all three authors. AC performed the analyses in Research Questions 2 and 3, and drafted the initial version of the paper, which was revised iteratively and approved by all three authors.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Inter-rater reliability.

[PDF File (Adobe PDF File), 45KB - jmir_v19i9e292_app1.pdf]

Multimedia Appendix 2

Tags and their assigned categories.

[CSV File, 13KB - jmir_v19i9e292_app2.csv]

References

1. Gruber T. Collective knowledge systems: where the social web meets the semantic web. Web Semant 2008 Feb;6(1):4-13. [doi: 10.1016/j.websem.2007.11.011]
2. Gupta M, Li R, Yin Z, Han J. Survey on social tagging techniques. ACM SIGKDD Explorations Newsletter 2010 Jun;12(1):58-72. [doi: 10.1145/1882471.1882480]
3. Hunter J, Gerber A. Harvesting community annotations on 3D models of museum artefacts to enhance knowledge, discovery and re-use. J Cult Herit 2010 Jan;11(1):81-90. [doi: 10.1016/j.culher.2009.04.004]
4. Sigurbjörnsson B, Van Zwol R. Flickr tag recommendation based on collective knowledge. New York, NY: ACM; 2008 Apr 21 Presented at: Proceedings of the 17th International Conference on World Wide Web; 2008; Beijing, China p. 327-326 URL: http://dl.acm.org/citation.cfm?id=1367542 [doi: 10.1145/1367497.1367542]
5. Yi K. Harnessing collective intelligence in social tagging using Delicious. J Am Soc Inf Sci Tec 2012 Nov 20;63(12):2488-2502. [doi: 10.1002/asi.22734]
6. Golder S, Huberman B. Usage patterns of collaborative tagging systems. J Inf Sci 2016 Jul;32(2):198-208. [doi: 10.1177/0165551506062337]

7. Syn S, Spring M. Finding subject terms for classificatory metadata from user-generated social tags. J Am Soc Inf Sci Tec 2013 Mar 14;64(5):964-980. [doi: 10.1002/asi.22804]
8. Choi Y. A complete assessment of tagging quality: a consolidated methodology. J Assn Inf Sci Tec 2014 May 15;66(4):798-817. [doi: 10.1002/asi.23198]
9. Macgregor G, McCulloch E. Collaborative tagging as a knowledge organisation and resource discovery tool. Libr Rev 2006 Jun;55(5):291-300. [doi: 10.1108/00242530610667558]
10. Lee DH, Schleyer T. Social tagging is no substitute for controlled indexing: a comparison of medical subject headings and CiteULike tags assigned to 231,388 papers. J Am Soc Inf Sci Tec 2012 Aug 17;63(9):1747-1757. [doi: 10.1002/asi.22653]
11. Mai J. Folksonomies and the new order: authority in the digital disorder. Knowl Organ 2011;38(2):114-122.
12. Trant J. Studying social tagging and folksonomy: A review and framework. Journal of Digital Information 2009;10(1):1-44 [FREE Full text]
13. Schuyler PL, Hole WT, Tuttle MS, Sherertz DD. The UMLS Metathesaurus: representing different views of biomedical concepts. Bull Med Libr Assoc 1993 Apr;81(2):217-222 [FREE Full text] [Medline: 8472007]
14. Lee D, de Keizer N, Lau F, Cornet R. Literature review of SNOMED CT use. J Am Med Inform Assoc 2014 Feb;21(e1):e11-e19 [FREE Full text] [doi: 10.1136/amiajnl-2013-001636] [Medline: 23828173]
15. Hartzler A, Pratt W. Managing the personal side of health: how patient expertise differs from the expertise of clinicians. J Med Internet Res 2011 Aug 16;13(3):e62 [FREE Full text] [doi: 10.2196/jmir.1728] [Medline: 21846635]
16. Doing-Harris KM, Zeng-Treitler Q. Computer-assisted update of a consumer health vocabulary through mining of social network data. J Med Internet Res 2011 May 17;13(2):e37 [FREE Full text] [doi: 10.2196/jmir.1636] [Medline: 21586386]
17. Vydiswaran V, Mei Q, Hanauer D, Zheng K. Mining consumer health vocabulary from community-generated text. AMIA Annu Symp Proc 2014;2014:1150-1159.
18. Smith CA, Wicks PJ. PatientsLikeMe: consumer health vocabulary as a folksonomy. AMIA Annu Symp Proc 2008 Nov 06:682-686 [FREE Full text] [Medline: 18999004]
19. Konstantinidis S, Fernandez-Luque L, Bamidis P, Karlsen R. The role of taxonomies in social media and the semantic web for health education. A study of SNOMED CT terms in YouTube health video tags. Methods Inf Med 2013;52(2):168-179. [doi: 10.3414/ME12-02-0005] [Medline: 23450378]
20. Dixon BE, McGowan JJ. Enhancing a taxonomy for health information technology: an exploratory study of user input towards folksonomy. Stud Health Technol Inform 2010;160(Pt 2):1055-1059. [Medline: 20841845]
21. Schulman-Green D, Jaser S, Martin F, Alonzo A, Grey M, McCorkle R, et al. Processes of self-management in chronic illness. J Nurs Scholarsh 2012 Jun;44(2):136-144 [FREE Full text] [doi: 10.1111/j.1547-5069.2012.01444.x] [Medline: 22551013]
22. Audulv Å. The over time development of chronic illness self-management patterns: a longitudinal qualitative study. BMC Public Health 2013 May 07;13:452 [FREE Full text] [doi: 10.1186/1471-2458-13-452] [Medline: 23647658]
23. Mehling WE, Gopisetty V, Daubenmier J, Price CJ, Hecht FM, Stewart A. Body awareness: construct and self-report measures. PLoS One 2009;4(5):e5614 [FREE Full text] [doi: 10.1371/journal.pone.0005614] [Medline: 19440300]
24. Chen AT. The relationship between health management and information behavior over time: a study of the illness journeys of people living with fibromyalgia. J Med Internet Res 2016 Oct 25;18(10):e269 [FREE Full text] [doi: 10.2196/jmir.5309] [Medline: 27780794]
25. Wolf CT, Veinot TC. Struggling for space and finding my place: An interactionist perspective on everyday use of biomedical information. J Assn Inf Sci Tec 2014 May 15;66(2):282-296. [doi: 10.1002/asi.23178]
26. Ehrlich C, St John W, Kendall E. 'Listening to my body' to 'Look after my body': a theory of information use for self-management of chronic obstructive pulmonary disease. J Nurs Healthc Chronic Illn 2010 Dec;2(4):262-270. [doi: 10.1111/j.1752-9824.2010.01067.x]
27. Hernandez CA, Bradish GI, Rodger NW, Rybansky SI. Self-awareness in diabetes: using body cues, circumstances, and strategies. Diabetes Educ 1999;25(4):576-584. [doi: 10.1177/014572179902500410] [Medline: 10614262]
28. Wilde M, Garvin S. A concept analysis of self-monitoring. J Adv Nurs 2007 Feb;57(3):339-350. [doi: 10.1111/j.1365-2648.2006.04089.x] [Medline: 17233653]
29. Bhimani RH, Gaugler JE, Skay C. Understanding symptom experiences of muscle tightness from patients' and clinicians' perspectives. J Clin Nurs 2017 Jul;26(13-14):1927-1938. [doi: 10.1111/jocn.13506] [Medline: 27533094]
30. Chen AT, Carriere R, Kaplan SJ, Colht K, Morey OT, Flaherty M, et al. Exploring collective tagging as a mechanism to elicit language about health management. In: Proc Assoc Info Sci Tech. 2016 Dec 27 Presented at: Annual Meeting of the Association for Information Science and Technology; October 14-18, 2016; Copenhagen, Denmark p. 1-6 URL: http://dl.acm.org/citation.cfm?id=3017447.3017551 [doi: 10.1002/pra2.2016.14505301104]
31. Aronson A, Lang FM. An overview of MetaMap: historical perspective and recent advances. J Am Med Inform Assoc 2010;17(3):229-236 [FREE Full text] [doi: 10.1136/jamia.2009.002733] [Medline: 20442139]
32. Creswick N, Westbrook JI, Braithwaite J. Understanding communication networks in the emergency department. BMC Health Serv Res 2009;9:247 [FREE Full text] [doi: 10.1186/1472-6963-9-247] [Medline: 20043845]

33. Gardy JL, Johnston JC, Ho Sui SJ, Cook VJ, Shah L, Brodkin E, et al. Whole-genome sequencing and social-network analysis of a tuberculosis outbreak. N Engl J Med 2011 Feb 24;364(8):730-739 [FREE Full text] [doi: 10.1056/NEJMoa1003176] [Medline: 21345102]
34. Zhang J, Xie J, Hou W, Tu X, Xu J, Song F, et al. Mapping the knowledge structure of research on patient adherence: knowledge domain visualization based co-word analysis and social network analysis. PLoS One 2012;7(4):e34497 [FREE Full text] [doi: 10.1371/journal.pone.0034497] [Medline: 22496819]
35. Wasserman S, Faust K. Social network analysis: methods and applications. Cambridge, UK: Cambridge University Press; 1994.
36. Newman MEJ. Modularity and community structure in networks. Proc Natl Acad Sci USA 2006 Jun 06;103(23):8577-8582 [FREE Full text] [doi: 10.1073/pnas.0601602103] [Medline: 16723398]
37. Steyvers M, Tenenbaum JB. The large-scale structure of semantic networks: statistical analyses and a model of semantic growth. Cogn Sci 2005 Jan 02;29(1):41-78. [doi: 10.1207/s15516709cog2901_3]
38. Bastian M, Heymann S, Jacomy M. Gephi: an open source software for exploring and manipulating networks. In: Proceedings of the Third International ICWSM Conference (2009). Palo Alto, CA: AAAI Publications; 2009 Presented at: Third International AAAI Conference on Weblogs and Social Media; May 17-20, 2009; San Jose, CA p. 362 URL: https://www.aaai.org/ocs/index.php/ICWSM/09/paper/view/154/1009
39. Jacomy M, Venturini T, Heymann S, Bastian M. ForceAtlas2, a continuous graph layout algorithm for handy network visualization designed for the Gephi software. PLoS One 2014;9(6):e98679 [FREE Full text] [doi: 10.1371/journal.pone.0098679] [Medline: 24914678]
40. Nonnecke B, Preece J. Lurker demographics: counting the silent. In: Proceedings of the SIGCHI conference on Human Factors in Computing Systems. New York, NY: ACM; 2000 Presented at: SIGCHI conference on Human Factors in Computing Systems; April 01-06, 2000; The Hague, The Netherlands p. 73-80 URL: http://dl.acm.org/citation.cfm?id=332409 [doi: 10.1145/332040.332409]
41. Keselman A, Smith CA, Divita G, Kim H, Browne AC, Leroy G, et al. Consumer health concepts that do not map to the UMLS: where do they fit? J Am Med Inform Assoc 2008;15(4):496-505 [FREE Full text] [doi: 10.1197/jamia.M2599] [Medline: 18436906]
42. Ley T, Seitlinger P. Dynamics of human categorization in a collaborative tagging system: how social processes of semantic stabilization shape individual sensemaking. Comput Human Behav 2015 Oct;51(Pt A):140-151 [FREE Full text] [doi: 10.1016/j.chb.2015.04.053] [Medline: 26566299]
43. Batch Y, Yusof MM, Noah SA. ICDTag: a prototype for a web-based system for organizing physician-written blog posts using a hybrid taxonomy-folksonomy approach. J Med Internet Res 2013 Feb 27;15(2):e41 [FREE Full text] [doi: 10.2196/jmir.2353] [Medline: 23470419]
44. Kiu C, Tsui E. TaxoFolk: A hybrid taxonomy–folksonomy structure for knowledge classification and navigation. Expert Syst Appl 2011 May;38(5):6049-6058. [doi: 10.1016/j.eswa.2010.11.014]
45. Belém F, Almeida J, Gonçalves M. A survey on tag recommendation methods. J Assn Inf Sci Tec 2016 Dec 21;68(4):830-844. [doi: 10.1002/asi.23736]
46. Cantador I, Konstas I, Jose JM. Categorising social tags to improve folksonomy-based recommendations. Web Semant 2011 Mar;9(1):1-15. [doi: 10.1016/j.websem.2010.10.001]
47. Cattuto C, Benz D, Hotho A, Stumme G. Semantic grounding of tag relatedness in social bookmarking systems. : Springer; 2011 Mar Presented at: International Conference on The Semantic Web (ISWC '08); October 26-30, 2008; Karlsruhe, Germany p. 1-15 URL: http://link.springer.com/chapter/10.1007/978-3-540-88564-1_39 [doi: 10.1007/978-3-540-88564-1_39]
48. Laniado D, Eynard D, Colombetti M. Using WordNet to turn a folksonomy into a hierarchy of concepts. In: Semantic Web Application and Perspectives - Fourth Italian Semantic Web Workshop (SWAP 2007). 2007 Dec 18 Presented at: Semantic Web Application and Perspectives - Fourth Italian Semantic Web Workshop (SWAP 2007); December 18-20, 2007; Bari, Italy p. 192-201.
49. Körner C, Benz D, Hotho A, Strohmaier M, Stumme G. Stop thinking, start tagging: tag semantics emerge from collaborative verbosity. New York, NY: ACM; 2010 Apr 26 Presented at: Proceedings of the 19th international conference on world wide web (WWW '10); April 26-30, 2010; Raleigh, NC URL: http://dl.acm.org/citation.cfm?id=1772744 [doi: 10.1145/1772690.1772744]

Edited by G Eysenbach; submitted 18.03.17; peer-reviewed by S Adams, C Deccache, T Rasekaba; comments to author 03.05.17; revised version received 10.06.17; accepted 11.06.17; published 07.09.17.

Please cite as:
Chen AT, Carriere RM, Kaplan SJ
The User Knows What to Call It: Incorporating Patient Voice Through User-Contributed Tags on a Participatory Platform About Health Management
J Med Internet Res 2017;19(9):e292
URL: http://www.jmir.org/2017/9/e292/
doi:10.2196/jmir.7673
PMID:28882809

Original Paper

Evaluating Patient Empowerment in Association With eHealth Technology: Scoping Review

Tracie Risling[1*], RN, PhD; Juan Martinez[1*], BSc; Jeremy Young[1*], BSc; Nancy Thorp-Froslie[1*], BS, MPH

College of Nursing, University of Saskatchewan, Saskatoon, SK, Canada
[*]all authors contributed equally

Corresponding Author:
Tracie Risling, RN, PhD
College of Nursing
University of Saskatchewan
4218 Health Sciences Building, E-Wing
104 Clinic Place
Saskatoon, SK, S7N 2Z4
Canada
Phone: 1 3069666232
Email: tracie.risling@usask.ca

Abstract

Background: The prioritization of sustainable patient-centered care in contemporary health care has resulted in an increased focus on patient empowerment, which in turn is considered to facilitate patient independence, self-management, and self-efficacy. However, a definitional consensus of empowerment remains elusive, impeding efforts to translate the conceptual ideals of empowerment into a measurable entity associated with changes in health care behavior or outcomes. The rapid integration of technology in health care serves to add another layer of complexity in the measurability and operationalization of empowerment and helps to create a specific context in which this conceptual entity should be further examined.

Objective: The primary objective of this scoping review was to explore the concept of patient empowerment within the electronic health (eHealth) context. A further focus on the association or measurement of this concept in conjunction with tethered patient portal use was also employed.

Methods: In this scoping review, a six-step framework was used to guide the search and paper selection process. The review was initiated with two broad research questions, which are as follows: (1) What is the relationship between empowerment and the use of eHealth technologies from a patient perspective? (2) How is patient empowerment (and/or engagement or activation) influenced by accessing personal health information through a tethered patient portal? Multiple databases were employed in a comprehensive search strategy, and papers were primarily evaluated and selected for inclusion by 2 review authors, and a third author was consulted to resolve any issues in reaching consensus.

Results: From an initial count of 1387 publications, this review returned nine systematic or literature review papers and 19 empirical studies that pertained to patient empowerment (and/or engagement and activation) in relation to the use of tethered patient portals providing access to electronic health records (EHRs). Of the 19 empirical publications, only four were found to have used specific patient empowerment measures with significant variety in their identified conceptual elements.

Conclusions: There is a persistent lack of conceptual clarity in patient empowerment research, and this has extended to study within the eHealth context. The interchangeable use or conflation of terms such as patient empowerment, engagement, and activation, has further complicated the advancement of distinct conceptual measures. To more strongly align changes in patient empowerment with supportive eHealth solutions, the challenges of achieving a consensus on how best to operationalize and measure patient empowerment must be met.

(J Med Internet Res 2017;19(9):e329) doi:10.2196/jmir.7809

KEYWORDS

review; eHealth; patient engagement; patient empowerment; patient activation; measure

Introduction

In addition to complex technological evolution and advancements, health care systems are undergoing a significant paradigmatic shift in response to the demand for care transformations that deliver on the long-standing promise of patient-centered care. The move from a decidedly more paternalist system, dominated by the views and preferences of health care practitioners, to one in which patient voice has arisen as a priority, has resulted in an increased exploration of patient empowerment [1,2]. Conceptually appealing in numerous health care applications and explorations, patient empowerment is emerging as a focal point in health care research and reform [2]. Empowerment is considered to facilitate patient control through self-management and shared decision making, as well as promote equitable and collaborative approaches to health care and improved cost-effectiveness of care delivery [3]. However, significant challenges remain for those wanting to translate the conceptual ideal of patient empowerment into measurable changes in health care behaviors or outcomes.

One of the most persistent issues in the consistent operationalization of patient empowerment is a lack of a clear definition of this complex concept [1,4,5]. The multitude of applications of the term in the literature has established empowerment as a process, often of a transformative nature; a representation, or manifestation of purported key elements such as self-management and freedom of choice; an aspect or result of particular interventions themselves where often these results are measured through improved patient outcomes or reported self-management [1]. Whereas this conceptual manipulation of patient empowerment has supported a diverse array of associated study, it has been less helpful in the establishment of a concrete and comprehensive singular measure of patient empowerment.

In the pursuit of patient-centered care, this achievement of definitional consensus, a necessity to facilitate the consistent operationalization and subsequent measurement of patient empowerment, has so far remained elusive [4,5]. In addition to the broad use of the term itself, patient empowerment has also been used interchangeably with the terms patient engagement, patient enablement, patient activation, and even patient-centeredness, though numerous reports support the distinct use and application of each of these key conceptual entities [5-7]. The role of patient empowerment has been explored in specific care contexts [8], with particular chronic diagnoses such as diabetes [9,10] and cardiac conditions [11], and for patient populations spanning the full range from pediatric to geriatric. The diverse application of this popular concept has also extended into the electronic health (eHealth) literature [12-15], which, while beginning to explore important considerations regarding the influence of technology on empowerment, has also been challenged by this persistent conceptual conflation.

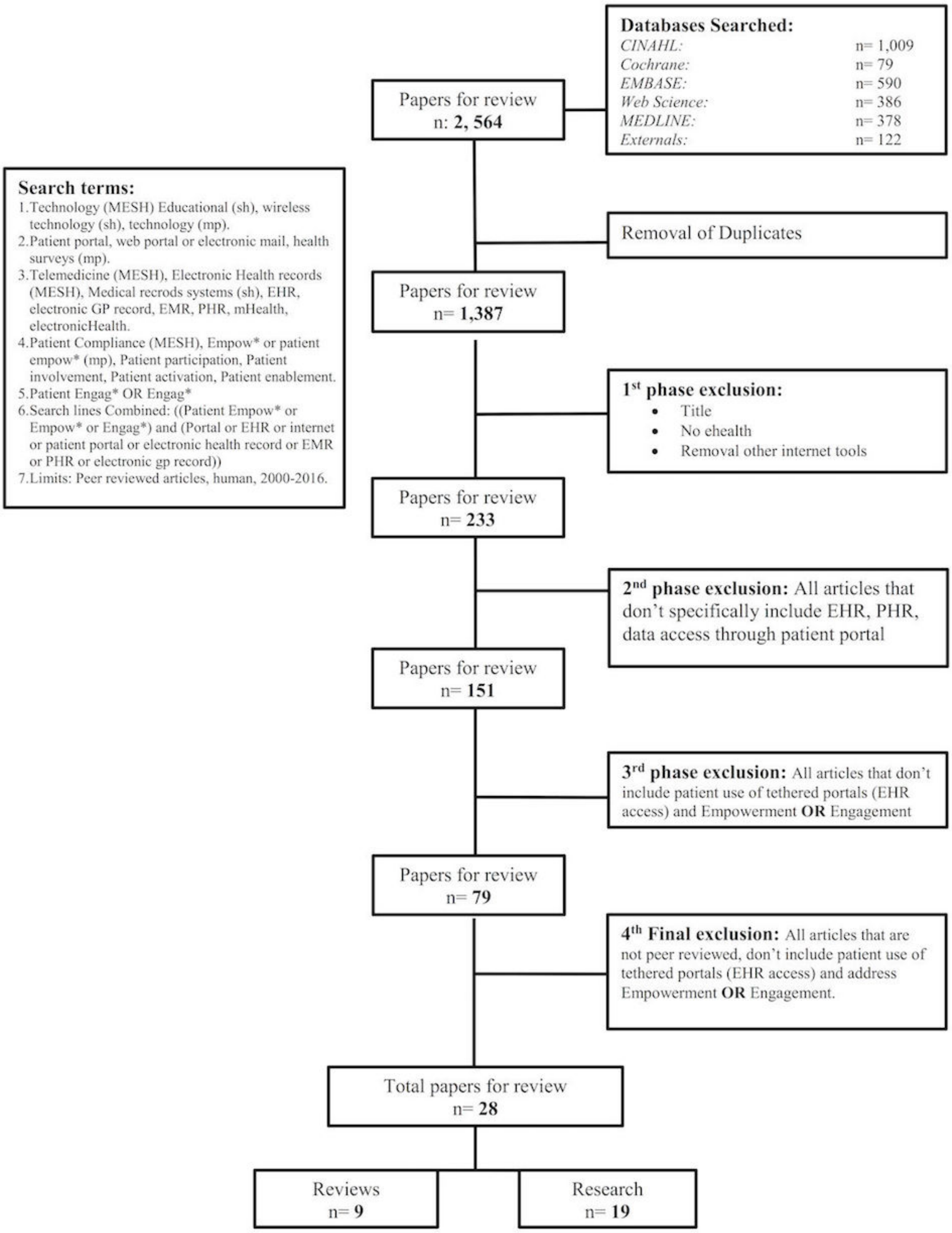
Discussion on the use of technology to advance patient empowerment [12,13,16,17] has taken into consideration how the concept may need to be reimagined within the eHealth context [15,18]. In addition, particular technologies such as patient portals have become a focal point in this research, with connections made between portal use, patient empowerment, engagement, and/or activation and ultimately, improved personal health outcomes [19-22]. This early work is a promising beginning in the exploration of eHealth and patient empowerment. However, the challenges of achieving a unified conceptual view of empowerment and perhaps more importantly, a single comprehensive empirical tool to evaluate empowerment in association with eHealth inventions remains. It is not sufficient to promote the empowering effects of new technologies without an accompanying evaluation of the actual influence of the intervention in this area. To advance this work, researchers need reliable measures of patient empowerment suitable for application in the eHealth context.

The initial primary focus of this scoping review was patient empowerment within the eHealth context, with a particular examination of the concept in association with patient portal use. Previously noted challenges regarding the interchangeable use of the terms patient empowerment, patient engagement, and patient activation were discovered in the early stages of the review process, and ultimately, resulted in an expansion of the search parameters. Owing to the lack of definitional consensus and consistent application of these concepts, additional search terms related to patient activation and engagement were incorporated to provide a more complete assessment of the current state of empirical patient empowerment measure in association with patient portal use. This paper includes the review findings and accompanying analysis summarized as follows: (1) characterizations of reported effects of portal use on patient empowerment, (2) identification of the range of patient empowerment, engagement, and activation measures reported in association with portal use, and (3) enumeration of differences in patient empowerment definition and measure.

Methods

This study was conducted based on the guidelines outlined by Levac et al [23] in their update to the work of Arksey and O’Malley in 2005. There are six steps to the review process in this framework, which are as follows: (1) identifying the research question, (2) identifying the relevant studies, (3) study selection, (4) charting the data, (5) collating, summarizing, and reporting results, and (6) consultation (optional) [23]. With respect to research questions, the recommendation in this framework is that scoping reviews employ both a broad and more focused question to provide a direction for the initial identification and eventual selection of relevant studies [23]. The broad directive in this review was as follows: (1) *What is the relationship between empowerment and the use of eHealth technologies from a patient perspective?* This question supported the primary focus of the review on patient empowerment within eHealth. The second research question provided additional parameters for the search and selection of publication by focusing the review on a particular eHealth solution. This question was also expanded to include additional terms, shortly after the review began, as has been detailed. The second question is as follows: (2) *How is patient empowerment (and/or engagement or activation) influenced by accessing personal health information through a tethered patient portal?* Tethered patient portals typically provide patients with access to information contained in their electronic health record (EHR), as opposed to personal health records (PHRs) which may not. Together, these questions directed the subsequent search and selection of relevant eHealth publications presented in these review findings.

Figure 1. Scoping review strategy and results.



Search Strategy

A preliminary search strategy for each selected database (Cumulative Index to Nursing and Allied Health Literature, The Cochrane Library, EMBASE, MEDLINE [Ovid], and Web of Science) was used to determine the Medical Subject Headings (MeSH) and keywords. The portal-related search terms were separated into three categories: empowerment, engagement / activation, and technology. MeSH Headings and text words were removed from the search strategy where no database results were returned. A full overview of this process is included in the detailed search strategy depicted in Figure 1. The search was restricted to English language publications between 2006 and 2016. An additional manual search of abstracts was performed to complete the review and support the inclusion of relevant publication that was not returned by the search strategy primarily because of indexing error. This work addressed the second step in the review process, identifying relevant studies, and produced a large initial publication count (n=1387) for review.

Study Selection

In the third step of the review process, inclusion and exclusion criteria were used on the basis of the specifics of the stated research questions. These criteria can evolve, as was the case in this review, as the researchers became more familiar with the subject material during the course of the ongoing review of the publications [23]. Following the removal of duplicates obtained in the initial search, this review progressed through four phases of paper exclusion, as was detailed in the search strategy in Figure 1. The first and second authors did the majority of the exclusion, with the consultation of a third party, as needed, when consensus on exclusion could not be reached. A title review was done as a part of the first exclusion phase to remove papers that did not have a specific eHealth focus. This aided in the management of a large number of papers returned by the use of technology as a search term and resulted in a sample size (n=233) for further examination. In the second phase, an abstract review assisted the authors in removing papers that did not include studies on EHRs, PHRs, or patient portals (n=151). The third phase of the exclusion focused on identifying tethered portal publication that also included the concept of patient empowerment or patient engagement or activation (n=79). Finally, in a fourth exclusion, the authors reviewed full-text documents and ensured that each included study reported on tethered patient portals that provided patients access to their EHR and addressed patient empowerment and/or patient engagement or activation.

The final count of papers for the literature review was settled at 28, including nine literature or systematic review publications and 19 empirical study publications. The review publication was retained because of challenges in trying to delineate a clear distinction between patient empowerment, engagement, activation, and other terms, considering that it could prove to be valuable for the discussion of empirical findings.

Data Extraction

The empirical studies were screened for data, and key points were extracted and summarized in the Multimedia Appendices 1 and Table 1 that follow. In addition to collating characteristics such as sample size, country of origin, research approach and design (Multimedia Appendix 1), a specific examination on reported tools and measures was also completed (Multimedia Appendix 2). During the synthesis process, the review of reported tools and measures led to a further enumeration of measures of patient activation, engagement, or empowerment (Table 1). This process addressed the fourth step in the scoping review framework and the concluding steps (collating, summarizing, and reporting the results, and consultation) as has been represented in the discussion of the review results.

Table 1. Summary of tools and concepts specifically used to measure patient empowerment, engagement, and activation.

Concept	Authors	Method or tool: concepts measured
Activation	Ancker et al [35]	Patient Activation Measure (PAM): patient knowledge, skill, confidence for self-care
	Crouch et al [28]	PAM
	O’Leary et al [36]	PAM
	Riippa et al [34]	PAM
Engagement	Toscos et al [24]	PAM
	Shi et al [37]	PAM
	Gee et al [25]	Qualitative: self-management
	Pillemer et al [40]	Qualitative: self-control, knowledge
	Rief et al [41]	Qualitative: knowledge of patient role, self-efficacy, initiative, and commitment to care
	Shade et al [38]	Outcome measure: use of health care services
	Henry et al [39]	Outcome measure: care gap closures
Empowerment	Crouch et al [28]	Healthcare Empowerment Inventory, based on Health Care Empowerment Model: engagement, informed, collaboration, commitment to treatment and tolerance of uncertainties of outcomes [32]
	van der Vaart et al [30]	Different scales for each component: patient satisfaction with care, trusting physician-patient relationship, self-efficacy in provider-patient communication, perception of illness and personal control, medication adherence
	Tuil et al [29]	Different scales for each component: self-efficacy, knowledge about treatment, involvement in decision process
	Earnest et al [31]	Composite empowerment scale: control of care, knowledge of condition, preparedness, reassurance, understanding of provider instructions, trust, ability to find mistakes

Results

Characterizing Effects of Portal Use on Patient Empowerment and Associated Concepts

The scoping review resulted in the extraction of nine systematic or literature reviews and 19 empirical studies. Multimedia Appendix 1 summarizes key characteristics of the 19 empirical studies. There is a wide range of literature available, characterizing the effect of patient portal utilization on patient empowerment, and/or patient engagement and activation. The progress of research publications in this area has been on an increase from 16 to the 19 publications in this area, since 2014. There is a range of international publications on this topic that is available with the majority of the studies originating in the United States.

Identifying a Range of Conceptual Measures

In general, this review comprised a range of measures applied to the evaluation of patient empowerment, engagement, and activation among portal users. These measures were often accompanied by other assessments pertaining to specific health outcomes, medication adherence, patient and provider attitudes, or patient satisfaction. In addition, several qualitative explorations reported on patient experiences, perceived barriers associated with portal use, and perceived sense of empowerment [24-27].

Enumeration of Differences in Conceptual Operationalization and Measures

Patient empowerment, engagement, and activation emerged as the primary conceptual entities in this review. There was considerable discussion on the barriers, experiences, characteristics of users, and health outcomes associated with the use of patient portals, however, a more elaborate exploration of each of these parameters is beyond the scope of this paper. Table 1 provides an inventory of the review papers, per concept, including the details of the measures used as reported in the publications.

The analysis of the empirical publication in this review revealed that patient empowerment, even when highlighted as a primary focus in a publication or study design, was not always specifically measured as a distinct concept in the resultant research. As summarized in Table 1, of the 19 publications reviewed, only four reported the use of focused patient empowerment measures [28-31]. When used, explicit measures of empowerment were conducted with specific composite empowerment scales [28,30], or independent measurements of interrelated components deemed to be related to empowerment [29,31]. Key components defined in relation to empowerment were distinct and varied depending on the study publication [28-31]. Crouch et al [28] utilized a previously established 8-item Healthcare Empowerment Inventory [32], designed to specifically measure empowerment through reports of being engaged, informed, collaborative, committed to treatment, and tolerant of uncertainty to outcomes or trajectory, a noteworthy contribution to this area of study.

The use and measure of patient activation was the most conceptually focused finding in this review. The widespread uptake of the Patient Activation Measure (PAM) by Hibbard Stockard [33] as a standard in this area has provided some conceptual stability for patient activation. The use of this measure as an extension to measure patient empowerment and engagement may be complicating the achievement of similar conceptual clarity for these entities. Several studies utilized the PAM in their evaluation of the potentially empowering or engaging effects of portals on patient activation [28,34-36]. In this grouping of PAM-based studies, two focused solely on activation and portal use [34,35], one on how activation and knowledge can influence engagement [36], and finally, a single study that used a separate empowerment measure in addition to the PAM [28]. This work demonstrates a focused pursuit of activation study, and in one case, acknowledges a distinction between activation and empowerment through the use of separate measures [28]. However, the issue remains that in many publications the words empowering or patient empowerment are employed to introduce or provide context for the research, with little or no conceptual follow-up or clarification found in the study results. The reported conceptualization of patient engagement was also found to be similarly inconsistent, with two studies utilizing the PAM as a validated proxy measure of engagement [24,37]. Other studies defined engagement in relation to the patients’ utilization of health care services [38], self-management of care gap closures [39], and perceived feelings of self-control and management over their health [25,40]. However, within these publications, there was a noted absence in discussing and clarifying the conceptual uniqueness of patient empowerment, engagement, and/or activation.

Discussion

Principal Findings

The results of this scoping review have been summarized as follows: (1) characterizations of reported effects of portal use on patient empowerment; (2) identification of the range of patient empowerment, engagement, and activation measures reported in association with portal use; and (3) enumeration of differences in the patient empowerment definition and measure. Further reflection on the conceptual complexity uncovered in this review, including findings from recent systematic and literature reviews have been detailed here.

This work has revealed the effects of tethered patient portal utilization on patient empowerment, engagement, and/or activation as somewhat controversial. Whereas none of the previous systematic reviews retained in this scoping study directly focused on the effect of eHealth technologies on patient empowerment [16,42-49], eHealth interventions, in general, have been hypothesized to contribute to patient empowerment by increasing self-efficacy and providing tools for self-management [12,29]. Overall, in early research and review, utilization of portals has most commonly been associated with small changes in patient empowerment or activation [28-31,35]. However, the use of portals was found to result in improved self-reported levels of engagement or activation related to self-management [40,41] and enhanced knowledge [25]. Portal use was also positively associated with better health outcomes in various study populations [10,24,28,37]. For example, diabetic patients with access to an electronic patient portal demonstrated improved glycemic control [10]. The diverse nature of these types of measures, and their potential relationship with any accompanying conceptual evaluation, adds to the challenge of attempting to isolate or demonstrate significant change in concepts such as patient empowerment.

The true influence of patient portals on empowerment seems to be obscured by the lack of a common vision of the concept itself. This is a significant consideration, as the conceptual clarification of empowerment is tantamount to the translation and operationalization of the concept into concrete use and practice [12]. Out of the 19 empirical studies selected from the scoping review, only four publications measured empowerment specifically, and within these, the conceptual elements of empowerment differed significantly as has been recorded in Table 1 [28-31]. Furthermore, while the *PAM* remains a standard for the measurement of patient activation [28,34-36] and engagement by proxy [24,37], specific measures of empowerment identified in the study publications varied from the use of composite empowerment scales [28,31] to the use of independent scales measuring interrelated components of empowerment [29,31] (Table 1). Hence, our review of the literature identified an overall lack of clarity and consensus surrounding the measurement and concepts related to patient empowerment in association with tethered patient portal usage.

Limitations

The perspective presented in this scoping review is limited to patient empowerment–focused research in relation to tethered patient portal use. As a consequence, a full exploration of patient empowerment research within a broader health care context was not completed, excluding studies which may lead to differing conclusions on the current state of patient empowerment measure. This said, measures utilized in more general health care application did present within this eHealth literature. Additionally, there were challenges in the applied search strategy on account of the use of ubiquitous terms such as Internet and empowerment. Although this was done to maximize a full scope of return and to combat noted indexing errors during the search, it also resulted in the return of many results not relevant to the focused questions that then had to be eliminated in the exclusion phases. The employed search strategy did not deliver a small number of key papers that were discovered during manual searches, revealing the potential indexing challenges. Ultimately, these publications were included as hand-searched items in the review to ensure a comprehensive body of literature from which to complete the exclusion process.

Comparison With Prior Work

This scoping review adds further evidence to prior publications that identified a lack of clarity and unification in the conceptualization of patient empowerment. There is substantial interest in achieving a more concrete operationalization of this concept [1,3,5,6]. The *Health Care Empowerment Inventory* [32] is one of the few empowerment focused measures in practice with others (not identified in this scoping review) such as the *Patient Empowerment Scale* [50], and a more recent tool, as yet unnamed, designed to evaluate patient empowerment in long-term conditions [51].The more recent work of Barr et al [5] has produced a comprehensive interdisciplinary conceptual map of empowerment, but these authors also noted that no existing single measure could adequately capture the complexity of the conceptual elements they had identified.

The findings of this review, in combination with prior studies, strengthens the position that patient empowerment is a distinct conceptual entity and should not be used interchangeably, with respect to terms such as engagement and activation [6,32]. However, numerous barriers to the development of a standardized measure of empowerment have been identified, such as the differing contexts and study populations in which the concept has been studied and diversified [12,52]; the prevalence of potentially conflicting or interrelated factors such as socioeconomic status, preexisting health conditions, Internet, or digital literacy [1]; possible issues with patient privacy and confidentiality [31]; and concerns regarding the use of technology in advancing patient empowerment, potentially resulting in a digital divide [28]. The interplay between patient empowerment, engagement, and activation must also be more clearly articulated, especially as it is related to the uptake of eHealth solutions. For example, it has been hypothesized that a high level of patient activation is required before effective portal use [35], this would, in theory, be essential for an improved sense of empowerment in relation to the solution. Barr et al [5] also indicated that advances in empowerment measure are impeded not only by an overlap of terminology in application, but by a lack of accompanying robust psychometric evaluative data.

Given the challenges in capturing and measuring the full scope of patient empowerment, it is not surprising to discover conflicting reports regarding the use of eHealth services on resultant health outcomes, as highlighted in this work, and this is consistent with the conclusions in several previous systematic eHealth reviews [42-45,47-49]. Whereas some study populations did report positive benefits from using eHealth technologies [25,27,40,41,46], it was argued that these perceived effects would not necessarily translate to overall improvements in patient empowerment or health outcomes [42,43]. This paper contributes to a growing body of eHealth research on patient empowerment. The ever-expanding presence of eHealth in the health care landscape must be factored into continuing patient empowerment study [12,13,15].

Future Directions

The concept of patient empowerment lies within the scope of eHealth literature [12-15], with an emerging consideration that “the future of patient empowerment may lie in technological advancements and better access of patients to these technologies” [11]. This review has identified ongoing challenges regarding conflation and inconsistent conceptual application in this field and further demonstrated a current lack of consensus surrounding the operationalization and measurement of patient empowerment in particular. In addition, there are further issues with the inconsistency present in identified patient empowerment measures.

A clear association between the use of eHealth solutions, patient empowerment, and health outcomes remains elusive. Further patient-driven investigation on patient empowerment is an urgent need, particularly within certain contexts, such as eHealth intervention, where there is a paucity of literature. Until a comprehensive measure of patient empowerment is developed and thoroughly evaluated, significant challenges will remain within the eHealth context with respect to establishing patient empowerment as a means to positively influence health outcomes.

Conclusions

The aforementioned challenges in evaluating patient empowerment have influenced the effectiveness of research on the relation of this concept with specific and measurable changes in health outcomes. Even so, recent reviews on patient empowerment reveal global interest in the advancement of research on this concept [1,5,8]. The World Health Organization (WHO) European Regional Office included empowerment and patient-centered practice as key elements in its Health 2020 report [53], a follow-up on previous WHO study on the effectiveness of empowerment to improve health [54]. The earlier WHO research identified empowerment as an essential public health strategy but also called for the ongoing refinement of measures to evaluate empowerment [54]; yet, calls for improved definition of patient empowerment and measures to comprehensively evaluate the concept remain [1,5,8], and this is clearly a need for eHealth science as well. Despite early review and research, a reliable and valid measure to evaluate patient empowerment remains elusive. “Patient empowerment strategies have been shown to positively impact health care outcomes and will likely help shape the future of medical practice” [11]; however, without adequate measure, the researchers, practitioners, and program providers will be challenged to establish the value of these interventions. This scoping review is part of a larger research project examining empowerment and patient portal use. The results of this review will be united with qualitative interview data from patient users of an EHR patient portal to produce a more comprehensive patient-directed view of empowerment to support ongoing examination of the significance of this concept in eHealth.

Acknowledgments

This scoping review was completed with research funds provided by eHealth Saskatchewan, as part of a study on the pilot launch of the Citizen Health Information Portal. The review was conducted independently by a research team at the University of Saskatchewan through the College of Nursing. The authors would also like to acknowledge the support of the University of Saskatchewan in providing funds for the undergraduate summer research assistant program.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Summary of characteristics of empirical studies included in the scoping review.

[PDF File (Adobe PDF File), 288KB - jmir_v19i9e329_app1.pdf]

Multimedia Appendix 2

Summary of concepts, tools, and measures included in empirical studies used to characterize patient outcomes related to use of patient portals.

[PDF File (Adobe PDF File), 277KB - jmir_v19i9e329_app2.pdf]

References

1. Bravo P, Edwards A, Barr PJ, Scholl I, Elwyn G, McAllister M, Cochrane Healthcare Quality Research Group, Cardiff University. Conceptualising patient empowerment: a mixed methods study. BMC Health Serv Res 2015 Jul 01;15:252 [FREE Full text] [doi: 10.1186/s12913-015-0907-z] [Medline: 26126998]
2. Snowdon A, Alessi C, Schnarr K. sites.ivey.ca. 2014. “It's all about me”: the personalization of health systems URL: http://sites.ivey.ca/healthinnovation/files/2014/02/Its-All-About-Me-The-Personalization-of-Health-Systems.pdf [accessed 2017-01-31] [WebCite Cache ID 6n3qZfWeH]
3. McAllister M, Dunn G, Payne K, Davies L, Todd C. Patient empowerment: the need to consider it as a measurable patient-reported outcome for chronic conditions. BMC Health Serv Res 2012 Jun 13;12:157 [FREE Full text] [doi: 10.1186/1472-6963-12-157] [Medline: 22694747]
4. Herbert RJ, Gagnon AJ, Rennick JE, O'Loughlin JL. A systematic review of questionnaires measuring health-related empowerment. Res Theory Nurs Pract 2009;23(2):107-132. [Medline: 19558027]
5. Barr PJ, Scholl I, Bravo P, Faber MJ, Elwyn G, McAllister M. Assessment of patient empowerment--a systematic review of measures. PLoS One 2015;10(5):e0126553 [FREE Full text] [doi: 10.1371/journal.pone.0126553] [Medline: 25970618]

6. Fumagalli LP, Radaelli G, Lettieri E, Bertele' P, Masella C. Patient empowerment and its neighbours: clarifying the boundaries and their mutual relationships. Health Policy 2015 Mar;119(3):384-394. [doi: 10.1016/j.healthpol.2014.10.017] [Medline: 25467286]
7. Holmström I, Röing M. The relation between patient-centeredness and patient empowerment: a discussion on concepts. Patient Educ Couns 2010 May;79(2):167-172. [doi: 10.1016/j.pec.2009.08.008] [Medline: 19748203]
8. Cerezo PG, Juvé-Udina ME, Delgado-Hito P. Concepts and measures of patient empowerment: a comprehensive review. Rev Esc Enferm USP 2016;50(4):667-674 [FREE Full text] [doi: 10.1590/S0080-623420160000500018] [Medline: 27680054]
9. Rossi MC, Lucisano G, Funnell M, Pintaudi B, Bulotta A, Gentile S, BENCH-D Study Group. Interplay among patient empowerment and clinical and person-centered outcomes in type 2 diabetes. The BENCH-D study. Patient Educ Couns 2015 Sep;98(9):1142-1149. [doi: 10.1016/j.pec.2015.05.012] [Medline: 26049679]
10. Lau M, Campbell H, Tang T, Thompson DJ, Elliott T. Impact of patient use of an online patient portal on diabetes outcomes. Can J Diabetes 2014 Feb;38(1):17-21. [doi: 10.1016/j.jcjd.2013.10.005] [Medline: 24485208]
11. Kambhampati S, Ashvetiya T, Stone NJ, Blumenthal RS, Martin SS. Shared decision-making and patient empowerment in preventive cardiology. Curr Cardiol Rep 2016 Apr 20;18(5):49. [doi: 10.1007/s11886-016-0729-6] [Medline: 27098670]
12. Calvillo J, Román I, Roa LM. How technology is empowering patients? A literature review. Health Expect 2015 Oct;18(5):643-652 [FREE Full text] [doi: 10.1111/hex.12089] [Medline: 23711169]
13. Lettieri E, Fumagalli LP, Radaelli G, Bertele' P, Vogt J, Hammerschmidt R, et al. Empowering patients through eHealth: a case report of a pan-European project. BMC Health Serv Res 2015 Aug 05;15:309 [FREE Full text] [doi: 10.1186/s12913-015-0983-0] [Medline: 26242863]
14. Hill R, Betts LR, Gardner SE. Older adults' experiences and perceptions of digital technology: (dis)empowerment, wellbeing, and inclusion. Comput Human Behav 2015 Jul;48:415-423. [doi: 10.1016/j.chb.2015.01.062]
15. Bos L, Marsh A, Carroll D, Gupta S, Rees M. Patient 2. 2008 Presented at: International Conference on Semantic Web & Web Services; 2008; Las Vegas p. 164-167 URL: http://www.p4f.be/P4F/_images/20100408patientempowermenthealth20.pdf
16. 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]
17. Lober WB, Flowers JL. Consumer empowerment in health care amid the internet and social media. Semin Oncol Nurs 2011 Aug;27(3):169-182. [doi: 10.1016/j.soncn.2011.04.002] [Medline: 21783008]
18. Bos L. Patient empowerment: a two Way road. In: Wickramasinghe N, Bali R, Suomi R, Kirn S, editors. Critical Issues for the Development of Sustainable E-Health Solutions. Boston, MA: Springer; 2012:203-227.
19. Cahill JE, Gilbert MR, Armstrong TS. Personal health records as portal to the electronic medical record. J Neurooncol 2014 Mar;117(1):1-6. [doi: 10.1007/s11060-013-1333-x] [Medline: 24477621]
20. Steele GC, Miller D, Kuluski K, Cott C. Tying eHealth tools to patient needs: exploring the use of eHealth for community-dwelling patients with complex chronic disease and disability. JMIR Res Protoc 2014 Nov 26;3(4):e67 [FREE Full text] [doi: 10.2196/resprot.3500] [Medline: 25428028]
21. Kannry J, Beuria P, Wang E, Nissim J. Personal health records: meaningful use, but for whom? Mt Sinai J Med 2012;79(5):593-602. [doi: 10.1002/msj.21334] [Medline: 22976365]
22. Chen C, Weider K, Konopka K, Danis M. Incorporation of socioeconomic status indicators into policies for the meaningful use of electronic health records. J Health Care Poor Underserved 2014 Feb;25(1):1-16 [FREE Full text] [doi: 10.1353/hpu.2014.0040] [Medline: 24509007]
23. 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]
24. Toscos T, Daley C, Heral L, Doshi R, Chen Y, Eckert GJ, et al. Impact of electronic personal health record use on engagement and intermediate health outcomes among cardiac patients: a quasi-experimental study. J Am Med Inform Assoc 2016 Jan;23(1):119-128. [doi: 10.1093/jamia/ocv164] [Medline: 26912538]
25. Gee PM, Paterniti DA, Ward D, Soederberg Miller LM. e-patients perceptions of using personal health records for self-management support of chronic illness. Comput Inform Nurs 2015 Jun;33(6):229-237. [doi: 10.1097/CIN.0000000000000151] [Medline: 25899440]
26. Shah SG, Fitton R, Hannan A, Fisher B, Young T, Barnett J. Accessing personal medical records online: a means to what ends? Int J Med Inform 2015 Feb;84(2):111-118 [FREE Full text] [doi: 10.1016/j.ijmedinf.2014.10.005] [Medline: 25453275]
27. Woods SS, Schwartz E, Tuepker A, Press NA, Nazi KM, Turvey CL, et al. Patient experiences with full electronic access to health records and clinical notes through the My HealtheVet Personal Health Record Pilot: qualitative study. J Med Internet Res 2013 Mar 27;15(3):e65 [FREE Full text] [doi: 10.2196/jmir.2356] [Medline: 23535584]
28. Crouch PC, Rose CD, Johnson M, Janson SL. A pilot study to evaluate the magnitude of association of the use of electronic personal health records with patient activation and empowerment in HIV-infected veterans. PeerJ 2015;3:e852 [FREE Full text] [doi: 10.7717/peerj.852] [Medline: 25802815]

29. Tuil WS, Verhaak CM, Braat DD, de Vries Robbé PF, Kremer JA. Empowering patients undergoing in vitro fertilization by providing Internet access to medical data. Fertil Steril 2007 Aug;88(2):361-368. [doi: 10.1016/j.fertnstert.2006.11.197] [Medline: 17416366]
30. van der Vaart R, Drossaert CH, Taal E, Drossaers-Bakker KW, Vonkeman HE, van de Laar MA. Impact of patient-accessible electronic medical records in rheumatology: use, satisfaction and effects on empowerment among patients. BMC Musculoskelet Disord 2014 Mar 26;15:102 [FREE Full text] [doi: 10.1186/1471-2474-15-102] [Medline: 24673997]
31. Earnest MA, Ross SE, Wittevrongel L, Moore LA, Lin C. Use of a patient-accessible electronic medical record in a practice for congestive heart failure: patient and physician experiences. J Am Med Inform Assoc 2004;11(5):410-417 [FREE Full text] [doi: 10.1197/jamia.M1479] [Medline: 15187074]
32. Johnson MO, Rose CD, Dilworth SE, Neilands TB. Advances in the conceptualization and measurement of Health Care Empowerment: development and validation of the Health Care Empowerment inventory. PLoS One 2012;7(9):e45692 [FREE Full text] [doi: 10.1371/journal.pone.0045692] [Medline: 23029184]
33. Hibbard JH, Stockard J, Mahoney ER, Tusler M. Development of the Patient Activation Measure (PAM): conceptualizing and measuring activation in patients and consumers. Health Serv Res 2004 Aug;39(4 Pt 1):1005-1026 [FREE Full text] [doi: 10.1111/j.1475-6773.2004.00269.x] [Medline: 15230939]
34. Riippa I, Linna M, Rönkkö I. The effect of a patient portal with electronic messaging on patient activation among chronically ill patients: controlled before-and-after study. J Med Internet Res 2014 Nov 19;16(11):e257 [FREE Full text] [doi: 10.2196/jmir.3462] [Medline: 25413368]
35. Ancker JS, Osorio SN, Cheriff A, Cole CL, Silver M, Kaushal R. Patient activation and use of an electronic patient portal. Inform Health Soc Care 2015;40(3):254-266. [doi: 10.3109/17538157.2014.908200] [Medline: 24786648]
36. O'Leary KJ, Lohman ME, Culver E, Killarney A, Randy Smith Jr G, Liebovitz DM. The effect of tablet computers with a mobile patient portal application on hospitalized patients' knowledge and activation. J Am Med Inform Assoc 2016 Jan;23(1):159-165. [doi: 10.1093/jamia/ocv058] [Medline: 26078412]
37. Shi Y, Fuentes-Caceres V, McHugh M, Greene J, Verevkina N, Casalino L, et al. Electronic health records and patient activation – their interactive role in medication adherence. In: Zheng X, Zeng DD, Chen H, Leischow SJ, editors. Smart Health: International Conference, ICSH 2015, Phoenix, AZ, USA, November 17-18, 2015. Revised Selected Papers. Switzerland: Springer; 2016:219-230.
38. Shade SB, Steward WT, Koester KA, Chakravarty D, Myers JJ. Health information technology interventions enhance care completion, engagement in HIV care and treatment, and viral suppression among HIV-infected patients in publicly funded settings. J Am Med Inform Assoc 2015 Apr;22(e1):e104-e111. [doi: 10.1136/amiajnl-2013-002623] [Medline: 25030033]
39. Henry SL, Shen E, Ahuja A, Gould MK, Kanter MH. The online personal action plan: a tool to transform patient-enabled preventive and chronic care. Am J Prev Med 2016 Jul;51(1):71-77. [doi: 10.1016/j.amepre.2015.11.014] [Medline: 26826751]
40. Pillemer F, Price RA, Paone S, Martich GD, Albert S, Haidari L, et al. Direct release of test results to patients increases patient engagement and utilization of care. PLoS One 2016;11(6):e0154743 [FREE Full text] [doi: 10.1371/journal.pone.0154743] [Medline: 27337092]
41. Rief JJ, Hamm ME, Zickmund SL, Nikolajski C, Lesky D, Hess R, et al. Using health information technology to foster engagement: patients' experiences with an active patient health record. Health Commun 2017 Mar;32(3):310-319. [doi: 10.1080/10410236.2016.1138378] [Medline: 27223684]
42. Ammenwerth E, Schnell-Inderst P, Hoerbst A. Patient empowerment by electronic health records: first results of a systematic review on the benefit of patient portals. Stud Health Technol Inform 2011;165:63-67. [doi: 10.3233/978-1-60750-735-2-63] [Medline: 21685587]
43. Ammenwerth E, Schnell-Inderst P, Hoerbst A. The impact of electronic patient portals on patient care: a systematic review of controlled trials. J Med Internet Res 2012 Nov 26;14(6):e162 [FREE Full text] [doi: 10.2196/jmir.2238] [Medline: 23183044]
44. 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]
45. Jilka SR, Callahan R, Sevdalis N, Mayer EK, Darzi A. “Nothing about me without me”: an interpretative review of patient accessible electronic health records. J Med Internet Res 2015 Jun 29;17(6):e161 [FREE Full text] [doi: 10.2196/jmir.4446] [Medline: 26123476]
46. Kruse CS, Argueta DA, Lopez L, Nair A. Patient and provider attitudes toward the use of patient portals for the management of chronic disease: a systematic review. J Med Internet Res 2015 Feb 20;17(2):e40 [FREE Full text] [doi: 10.2196/jmir.3703] [Medline: 25707035]
47. Kruse CS, Bolton K, Freriks G. The effect of patient portals on quality outcomes and its implications to meaningful use: a systematic review. J Med Internet Res 2015 Feb 10;17(2):e44 [FREE Full text] [doi: 10.2196/jmir.3171] [Medline: 25669240]
48. de Lusignan S, Mold F, Sheikh A, Majeed A, Wyatt JC, Quinn T, et al. Patients' online access to their electronic health records and linked online services: a systematic interpretative review. BMJ Open 2014 Sep 08;4(9):e006021 [FREE Full text] [doi: 10.1136/bmjopen-2014-006021] [Medline: 25200561]

49. Otte-Trojel T, de Bont A, Rundall TG, van de Klundert J. How outcomes are achieved through patient portals: a realist review. J Am Med Inform Assoc 2014;21(4):751-757 [FREE Full text] [doi: 10.1136/amiajnl-2013-002501] [Medline: 24503882]
50. Faulkner M. A measure of patient empowerment in hospital environments catering for older people. J Adv Nurs 2001 Jun 19;34(5):676-686. [doi: 10.1046/j.1365-2648.2001.01797.x] [Medline: 11380736]
51. Small N, Bower P, Chew-Graham CA, Whalley D, Protheroe J. Patient empowerment in long-term conditions: development and preliminary testing of a new measure. BMC Health Serv Res 2013 Jul 08;13:263 [FREE Full text] [doi: 10.1186/1472-6963-13-263] [Medline: 23835131]
52. Cerezo PG, Juvé-Udina ME, Delgado-Hito P. Concepts and measures of patient empowerment: a comprehensive review. Rev Esc Enferm USP 2016;50(4):667-674 [FREE Full text] [doi: 10.1590/S0080-623420160000500018] [Medline: 27680054]
53. World Health Organization. euro.who. Copenhagen: World Health Organization Regional Office for Europe; 2012. Health 2020: a European policy framework supporting action across government and society for health and well-being URL: http://www.euro.who.int/__data/assets/pdf_file/0009/169803/RC62wd09-Eng.pdf [accessed 2017-01-30] [WebCite Cache ID 6riDMJ5HB]
54. Wallerstein N. euro.who. Copenhagen: WHO Regional Office for Europe's Health Evidence Network (HEN); 2006. What is the evidence on effectiveness of empowerment to improve health? URL: http://www.euro.who.int/__data/assets/pdf_file/0010/74656/E88086.pdf [accessed 2017-01-31] [WebCite Cache ID 6riCvQHGr]

Abbreviations

eHealth: electronic health
EHR: electronic health record
MeSH: medical subject headings
PAM: Patient Activation Measure
PHR: personal health record
RCT: randomized controlled trial
WHO: World Health Organization

Edited by G Eysenbach; submitted 02.04.17; peer-reviewed by M McNamara, P Gee, H Witteman; comments to author 11.05.17; revised version received 06.07.17; accepted 16.08.17; published 29.09.17.

Please cite as:
Risling T, Martinez J, Young J, Thorp-Froslie N
Evaluating Patient Empowerment in Association With eHealth Technology: Scoping Review
J Med Internet Res 2017;19(9):e329
URL: http://www.jmir.org/2017/9/e329/
doi:10.2196/jmir.7809
PMID:28963090

Original Paper

Effectiveness of Telemonitoring in Obstetrics: Scoping Review

Dorien Lanssens[1,2], MSc, Mw; Thijs Vandenberk[1], MSc; Inge M Thijs[1,3], Dr-Ing; Lars Grieten[1], MSc; Wilfried Gyselaers[1,2,3], Prof Dr Med

[1]Mobile Health Unit, Facultiy of Medicine and Life Sciences, Hasselt University, Hasselt, Belgium
[2]Department of Gynaecology, Ziekenhuis Oost-Limburg, Genk, Belgium
[3]Department of Physiology, Hasselt University, Hasselt, Belgium

Corresponding Author:
Dorien Lanssens, MSc, Mw
Department of Gynaecology
Ziekenhuis Oost-Limburg
Schiepse Bos 6
Genk, 3600
Belgium
Phone: 32 89321557
Fax: 32 89321557
Email: dorien.lanssens@uhasselt.be

Abstract

Background: Despite reported positive results of telemonitoring effectiveness in various health care domains, this new technology is rarely used in prenatal care. A few isolated investigations were performed in the past years but with conflicting results.

Objective: The aim of this review was to (1) assess whether telemonitoring adds any substantial benefit to this patient population and (2) identify research gaps in this area to suggest goals for future research.

Methods: This review includes studies exploring the effectiveness of telemonitoring interventions for pregnant women reported in the English language. Due to the paucity of research in this area, all reports including uncontrolled nonrandomized and randomized controlled studies were selected.

Results: Fourteen studies, which performed their data collection from 1988 to 2010, met the inclusion criteria and were published from 1995 to present; four of the 14 published papers were multicenter randomized controlled trials (RCTs), five papers were single-center RCTs, three papers were retrospective studies, one paper was an observational study, and one paper was a qualitative study. Of the 14 papers, nine were available for a risk of bias assessment: three papers were classified as *low risk*, one as *medium risk*, and five as *high risk*. Furthermore, of those 14 papers, 13 focused on telemonitoring for maternal outcomes, and nine of the 14 papers focused on telemonitoring for fetal or neonatal outcomes. The studies reviewed report that telemonitoring can contribute to significant reductions in health care costs, (unscheduled) face-to-face visits, low neonatal birth weight, and admissions to the neonatal intensive care unit (NICU), as well as prolonged gestational age and improved feelings of maternal satisfaction when compared with a control group. When only studies with low risk of bias were taken into account, the added value of telemonitoring became less pronounced: the only added value of telemonitoring is for pregnant women who transmitted their uterine activity by telecommunication. They had significant prolonged pregnancy survivals, and the newborns were less likely to be of low birth weight or to be admitted to the NICU. Following these results, telemonitoring can only be recommended by pregnant women at risk for preterm delivery. It is however important to consider that these studies were published in the mid-90s, which limits their direct applicability given the current technologies and practice.

Conclusions: This review shows that telemonitoring can be tentatively recommended for pregnant women at risk for preterm delivery. More recent RCTs with a blinded protocol are needed to strengthen the level of evidence around this topic and to have an insight in the added value of the technologies that are available nowadays. In addition, studies investigating patient satisfaction and economic effects in relation to telemonitoring are suggested for future research.

(J Med Internet Res 2017;19(9):e327) doi:10.2196/jmir.7266

KEYWORDS

review; telemonitoring; obstetrics; maternal outcomes; fetal outcomes

Introduction

With more than 6 billion mobile phone subscribers worldwide, it is estimated that 75% of the world population has access to mobile communication. The number of devices with broadband capabilities has increased to more than 1 billion worldwide [1]. With more than 97,000 health-related mobile apps available and approximately 1000 new apps published every month, the potential to perform telemedicine exists [1]. Telemedicine is a relatively new approach (dating back to the early 1990s), which facilitates patients' management at home [2]. It can be broadly defined as the use of telecommunication technologies to assist in the transmission of medical information and services between health care providers and patients. The use of this two-way telecommunication technology, multimedia, and computer networks to deliver or enhance the delivery of health care is a growing trend internationally [3]. It has the potential to improve access to high-quality disease management, and telemonitoring, a subgroup of telemedicine, has developed rapidly over the past decade [4]. There are several types of telemonitoring, ranging from simple to complex. In the simplest model, a patient receives support from a health care professional over the telephone. The patient monitors his or her symptoms and reports this during a structured telephone call. Moving up the scale of complexity is patient-initiated electronic monitoring with the transfer of physiologic data and record of symptoms by telephone or a broadband Internet connection from the patient's home (ie, home telemonitoring) to the health care professional. On reviewing the data, the health care professional can contact the patient to request further information before making a decision about disease management. Finally, implanted monitoring devices transmit data wirelessly from the patient to a unit that is connected to a telephone or the Internet. Once again, if the data raise concern, the health care professional can contact the patient to request further information before making a decision about care [4].

A number of systematic reviews have evaluated the effectiveness of telemonitoring interventions for patients diagnosed with chronic cardiovascular disease (CVD), chronic obstructive pulmonary disease (COPD), and diabetes [2,5-8]. These reviews show mainly positive results and suggest that there is tentative evidence that telemonitoring may offer clinical benefit in these three domains. All-cause mortality and heart-related hospitalizations are reduced for patients with CVD compared with patients who received usual care [5,6]. Even primary care management of CVD can be enhanced by improving patient outcomes and reducing health-related costs [7]. Web-based remote monitoring for managing type 2 diabetes mellitus is also a viable approach for health care delivery and enhances patients' quality of life [8]. Finally, home telemonitoring in patients with COPD appears to have a positive effect in reducing respiratory exacerbations and hospitalizations and in improving quality of life: patients with COPD were generally satisfied with home telemonitoring and found the systems useful to help them manage their disease and improve health care provision [2,9]. With regard to fertility, a few papers on telemonitoring discussing self-operated endovaginal telemonitoring of the ovarian stimulation phase in in vitro fertilization or intracytoplasmic sperm injection are published. This technique leads to relevant clinical decisions; significantly higher satisfaction of patients and their partner; a higher feeling of empowerment, discretion, and more active partner participation; as well as a trend toward less stress versus a traditional monitored group [10-12]. Despite the mainly positive results in the various health care domains and the ability to perform telemonitoring because of the improvement of technology, telemonitoring is rarely used in prenatal care. A few independent investigations were performed in the last years, but a systematic review has not yet been accomplished. For this reason, a systematic review of all clinical trials evaluating telemonitoring in high-risk pregnancies was performed. First, the characteristics of the study will be described, and then the maternal and neonatal outcomes in telemonitoring group versus control group (CG) will be reported. We aim to (1) assess whether telemonitoring adds any substantial benefit in the pregnant women population and (2) identify research gaps in this area and thereby suggest topics for future research.

Methods

Search Strategy

The following databases were comprehensively searched in August 2016 by two independent researches: the Medical Literature Analysis and Retrieval System Online (MEDLINE), the Cumulative Index to Nursing and Allied Health Literature (CINAHL), Excerpta Medica Database (EMBASE), PubMed, Limo, and the Cochrane Library. The enumeration of selected relevant journals was manually screened, and the bibliographies of all retained papers were examined for relevant studies. A third reviewer resolved discrepancies in judgment and verified the completeness of the manuscript.

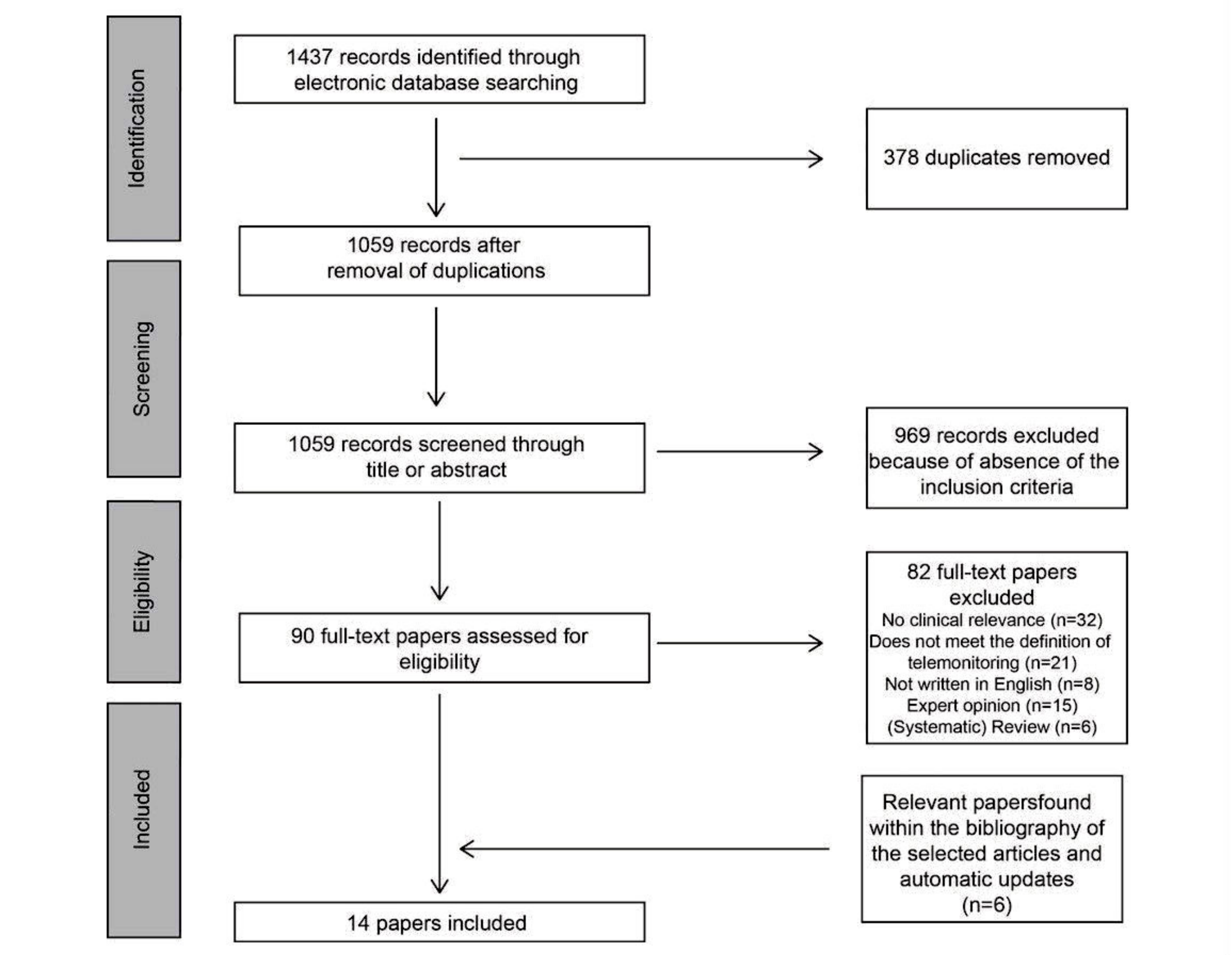
Search Items

The following terms were used in the search bar of the mentioned databases: "remote monitoring," "telemonitoring," "home monitoring," "telemedicine," "maternal health," "telehealth," "e-health," "pregnancy," "pregnancy-outcomes," "gynecology," "gravidity," and "obstetrics." Also medical subject headings (MeSH) thesaurus combined were used with the following terms: "blood pressure monitoring, ambulatory," "blood glucose self-monitoring," "pregnancy," pregnancy outcome," "gynecology," and "obstetrics."

A Definition of Telemonitoring

In this review, we specified the definition of telemonitoring—as stated in the introduction—further to the following inclusion criteria: (1) require the patient to periodically measure physiological parameters (eg, blood pressure and weight) and/or record their symptoms or vital signs in a standardized format, (2) use telecommunication technologies (eg, mobile phone and Internet) that either manually or automatically transferred the patient's health status data from home to a health care service, and (3) lead to the automated or manual review of the patient's health status data.

Figure 1. Selection procedure.



Inclusion and Exclusion Criteria

To be included, studies had to examine the effectiveness of telemonitoring interventions for pregnant women as defined above. Scoping searches indicated a paucity of research in this area, and we therefore included uncontrolled and nonrandomized, as well as randomized controlled studies. All published studies reporting economic and/or clinically related outcomes (eg, hospital admission and preterm labor) were considered. Due to the scarce available publications, no time limitations were applied. All papers had to be written in English. Studies were excluded if health care professionals conducted the measurement of physiological signs at the patient's home. In addition, review papers, expert opinions, and single case or case series reports were excluded.

Selection Procedure

A flowchart of the selection procedure is shown in Figure 1. The database search identified 1437 papers. After the removal of duplicates, 1059 records were screened for relevant content. During title, abstract, and keyword screening, 969 papers were excluded because of the absence of the inclusion criteria. The full-text of the 90 potentially relevant papers was assessed, and 82 papers were excluded. Reasons for exclusion included (1) no clinical or economical relevance (n=32), (2) does not meet the definition of telemonitoring (n=21), (3) not written in English (n=8), (4) expert opinions (n=15), and (5) (systematic) reviews (n=6). Automatic updates from the databases and search for relevant papers within the bibliography of selected papers retrieved six papers, which were also included. In total, 14 papers were included.

Assessment of Risk of Bias in Included Studies

A report on the methodological risk of bias of included studies (which had a randomized controlled design) in accordance with the Cochrane Handbook for Systematic Reviews of Interventions [13] and the guidelines of the Cochrane Consumers and Communication Review Group was made (Multimedia Appendix 1). These guidelines recommend the explicit reporting of the following individual elements for randomized controlled trials (RCTs): random sequence generation, allocation sequence concealment, blinding (participants and personnel), blinding (outcome assessment), completeness of outcome data, and selective outcome reporting. Each item is judged as being at high, low, or unclear risk of bias as set out in the criteria provided by Higgins et al (2011). Studies will be deemed to be at the highest risk of bias if they are scored as at high or unclear risk of bias for either the sequence generation or allocation concealment domains, based on growing empirical evidence that these factors are particularly important potential sources of bias [13].

Data Extraction

The following information was collected and tabulated from the included studies: description of patient population, sample size, whether any economic evaluation was performed, the nature of the intervention, and the outcomes reported.

Results

Study Characteristics

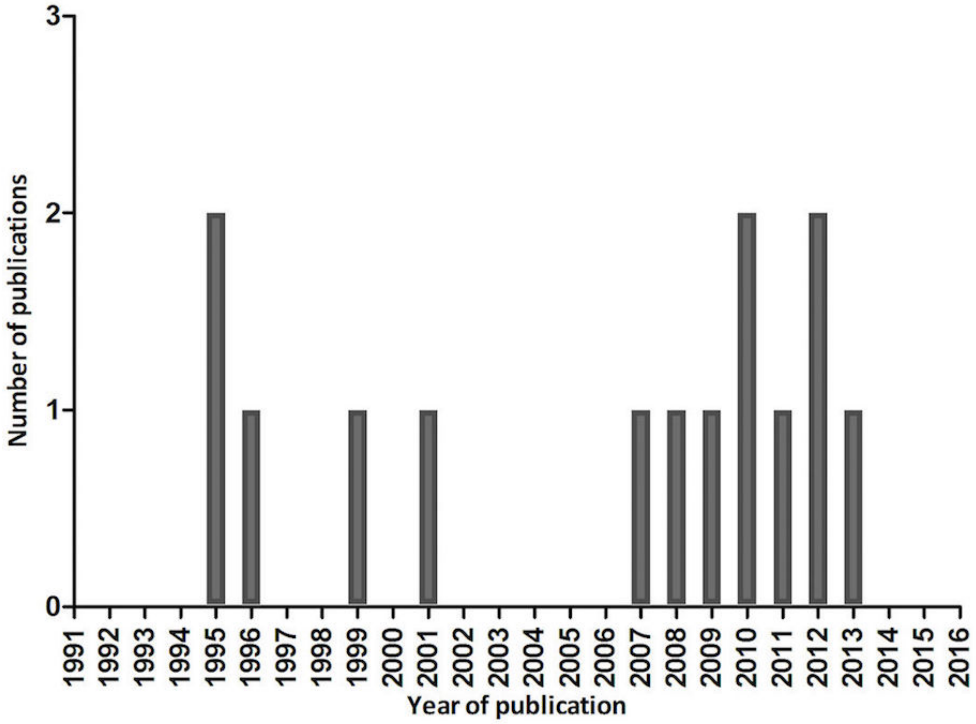
Fourteen studies were included, published from 1995 to present. An overview of these publication dates is presented in Figure 2.

Although the dates of the publications were from 1995 to present, the data collection was performed from 1988 to 2010.

Four of the 14 published papers were multicenter RCTs [14-17], five papers were single-center RCTs [18-22], three papers were retrospective studies [23-25], one paper was an observational study [26], and one paper was a qualitative study [27].

Multimedia Appendices 2 and 3 provide an overview of the characteristics of each study. All 14 papers report telemonitoring in obstetrics; 13 of the 14 papers focused on telemonitoring for maternal outcomes [14,15,17-27], and nine of the 14 papers focused on telemonitoring for fetal or neonatal outcomes [14-21,24]. Samples included varied from 15 singleton pregnancies [27] to 1292 singleton pregnancies [15]. Nine of the 14 papers were available for a risk of bias assessment (Multimedia Appendix 1): three papers were classified as *low risk* [14,15,25], one as *medium risk* [22], and five as *high risk* [17-21]. Five of the 14 papers did not have an RCT design [16,23,24,26,27]. For this reason, there was no risk of bias assessment made for them.

Figure 2. Number of publications during the last 25 years.



Telemonitoring data were generally transmitted to a monitoring center on a regular basis. Patients' compliance with data transmission was assessed in three studies [18,19,25] and ranged from a mean of 21.8 (standard deviation [SD] 16.9) sets of data [18] to a mean of 35.6 (SD 32.3) sets of data [19], depending on the physiological parameter measured. All the data were automatically transferred in the studies that investigated the added value of telemonitoring in pregnancies at high risk for preterm delivery or with an induction [14-16,23-26]. The data of the studies which investigated the added value of telemonitoring in gestational diabetes mellitus (GDM) were manually transferred [17-21]. In almost all the studies, patients' recordings outside predetermined values triggered an immediate action. Usual care included the same health care component as provided to the telemonitoring group but without telemonitoring.

Maternal Outcomes

Multimedia Appendix 2 provides a summary of the 12 studies included focusing on telemonitoring for maternal outcomes: cervical dilatation or preterm labor, GDM, maternal satisfaction, and health care–related costs. These results will be further discussed below.

Cervical Dilatation or Preterm Labor

The use of telemonitoring in the monitoring of fetal heart rate and uterine activity dates back to the 1970s. The expected benefits lie in the prevention of perinatal mortality and morbidity [3]. In five studies, women with singleton pregnancies at high risk for preterm birth were randomly assigned to a telemonitoring group and a CG. The results of these studies are presented in Table 1.

Gestational Diabetes Mellitus

The application of telemedicine in the management of GDM has primarily focused on the transfer of blood glucose values from the patient to the provider, thereby eliminating frequent clinical visits and adverse maternal and fetal or neonatal outcomes [3]. Five studies did report these study outcomes (Table 2).

Maternal Satisfaction

Due to the new aspect of telemonitoring, the maternal satisfaction of these domain is rarely investigated. Table 3 summarizes the major findings of five studies after adding telemonitoring to the obstetrical care.

Health Care–Related Costs

The continuous strain on hospital bed occupancy puts clinicians under great pressure to discharge patients as soon as possible. It is assumed that telemonitoring can contribute to solve this problem. Two studies did compute these costs for a telemonitoring group in comparison with a CG (Table 4).

Fetal or Neonatal Outcomes

Multimedia Appendix 3 provides a summary of the eight included studies focusing on telemonitoring for fetal or neonatal outcomes. In the next section, the influence of telemonitoring on the following fetal or neonatal outcomes will be presented: birth weight, gestational age, and submission to the NICU.

Birth Weight

Infants born small for gestational age (generally defined as less than 10th percentile) or large for gestational age (generally defined greater than 90th percentile) are at higher risk of short- and long-term morbidities than infants appropriately grown for gestational age [28]. A total of eight studies examined the impact of telemonitoring interventions on the birth weight of the neonate, which are presented in Table 5.

Table 1. Cervical dilatation or preterm labor and telemonitoring.

Citation	Risk of bias	Prolonged pregnancy survival	*P* value, TM[a] group versus CG[b]	Experience of a preterm delivery	*P* value, TM group versus CG
Brown et al (1999) [22]	Low risk			–[c]	.73
Corwin et al (1996) [14]	Low risk	+[d]	.02	–	.04
CHUMS[e] group (1995) [20]	Low risk			–	No[f,g]
Wapner et al (1995) [25]	Low risk	+	.016		
Morrison et al (2001) [24]	/[h]			–	<.001

[a] TM: telemonitoring.

[b] CG: control group.

[c] – indicates less experience in telemonitoring group versus control group.

[d] + indicates more experience in telemonitoring group versus control group.

[e] CHUMS: Collaborative Home Uterine Monitoring Study.

[f] NS: not significant.

[g] No exact value is given.

[h] The slash indicates there was no risk of bias assessment possible for the study because of study design (eg, observational study).

Table 2. Gestational diabetes mellitus and telemonitoring.

Citation	Risk of bias	Fasting blood sugar	*P* value, TM[a] group versus CG[b]	Glycated hemoglobin (HbA1c) <5.8%	*P* value, TM group versus CG	Insulin therapy	*P* value, TM group versus CG	Outpatient clinic visits	*P* value, TM group versus CG
Homko et al (2007) [18]	High risk	0[c]	No[d,e]			+[f]	<.05		
Dalfrà et al (2009) [17]	High risk			0	No	0	No[h]		
Homko et al (2012) [19]	High risk	0	.26			+	[e]		
Pérez-Ferre et al (2010) [20]	High risk							–[g]	<.001
Pérez-Ferre et al (2010) [21]	High risk			0	No			–	<.001

[a] TM: telemonitoring.

[b] CG: control group.

[c] 0: no differences.

[d] NS: not significant.

[e] no exact value is given.

[f] + indicates more experience in telemonitoring group versus CG.

[g] – indicates less experience in telemonitoring group versus CG.

Table 3. Maternal satisfaction and telemonitoring.

Citation	Risk of bias	Result for women in telemonitoring group
Homko et al (2007) [18]	High risk	More feelings of self-efficacy in women with GDM[a]
Dalfrà et al (2009) [17]	High risk	Women in the telemonitoring group showed lower levels of frustration and concerns about their GDM and a better acceptance of their diabetic condition
O'Brien et al (2013) [27]	/[b]	Better birth experiences resulting by induction of labor at home
Pérez-Ferre et al (2010) [21]	High risk	Higher patient satisfaction in women with GDM
Rauf et al (2011) [26]	/	Labor induction at home is feasible and acceptable to women

[a] GDM: gestational diabetes mellitus.

[b] The slash indicates there was no risk of bias assessment possible for the study because of study design (eg, observational study).

Table 4. Health care–related costs and telemonitoring.

Citation	Risk of bias	Result for women in telemonitoring group versus women in control group	
		Total cost saving	Average cost saving per pregnancy
Buysse et al (2008) [23]	/[a]	€145,822 for 415 pregnant women	€351.38
Morrison et al (2001) [24]	/	US $867,540 for 60 pregnant women	US $14,459

[a] The slash indicates there was no risk of bias assessment possible for the study because of study design (eg, observational study).

Table 5. Birth weight and telemonitoring.

Citation	Risk of bias	Small for gestational age (<10th percentile)	*P* value, TM[a] group versus CG[b]	Mean birth weight	*P* value, TM group versus CG	Large for gestational age (>90th percentile)	*P* value, TM group versus CG
CHUMS[c] group (1995) [20]	Low risk	–[d]	No[e,f]	+[g]	No[e]		
Corwin et al (1996) [14]	Low risk	–	.003				
Homko et al (2007) [18]	High risk					+	No[e]
Dalfrà et al (2009) [17]	High risk			0[h]	No[e]	0	No[e]
Homko et al (2012) [19]	High risk			0	.30	+	.70
Morrison et al (2001) [24]	/[i]	–	.001	+	<.001		
Pérez-Ferre et al (2010) [20]	High risk			0	No[e]		
Pérez-Ferre et al (2010) [21]	High risk			0	.39	-	No[e]

[a] TM: telemonitoring.

[b] CG: control group.

[c] CHUMS: Collaborative Home Uterine Monitoring Study.

[d] – indicates less experiences or lower mean in telemonitoring group versus CG.

[e] NS: not significant.

[f] No exact value is given.

[g] + indicates more experiences or higher mean in telemonitoring group versus CG.

[h] 0= no differences.

[i] The slash indicates there was no risk of bias assessment possible for the study because of study design (eg, observational study).

Table 6. Gestational age and telemonitoring.

Citation	Risk of bias	<37 weeks	*P* value, TM[a] group versus CG[b]	<36 weeks	*P* value, TM group versus CG	<35 weeks	*P* value, TM group versus CG	<34 weeks	*P* value, TM group versus CG	<32 weeks	*P* value, TM group versus CG
CHUMS[c] group (1995) [20]	Low risk	+[d]	No[e,f]	–	No[e]			–	No[f]		
Homko et al (2007) [18]	High risk	0	No[e]								
Morrison et al (2001) [24]	/[g]					–	<.01			–	.003
Kuleva et al (2012) [16]	/			–	.016						

[a] TM: telemonitoring.

[b] CG: control group.

[c] CHUMS: Collaborative Home Uterine Monitoring Study.

[d] + indicates more experiences or higher mean in telemonitoring group versus CG.

[e] NS: not significant.

[f] No exact value is given.

[g] The slash indicates there was no risk of bias assessment possible for the study because of study design (eg, observational study).

Gestational Age

We previously reported the influence of telemonitoring on cervical dilatation or preterm labor. One of the consequences of preterm labor is a preterm delivery of the newborn. Only four studies reported gestational age of the newborn as a main outcome. In Table 6, the rate of experiences of preterm births (for the gestational age of less than 37 weeks, less than 36 weeks, less than 35 weeks, less than 34 weeks, or less than 32 weeks) in telemonitoring group versus CG is reported.

Submission to Neonatal Intensive Care Unit (NICU)

Four studies have investigated the added value of telemonitoring and the submission to the NICU. These studies are presented in Table 7.

Table 7. Submission to the neonatal intensive care unit (NICU) and telemonitoring.

Citation	Risk of bias	Admission to NICU[a]	*P* value, telemonitoring group versus control group
CHUMS[b] group (1995) [20]	Low risk	–[c]	No[d,e]
Corwin et al (1996) [14]	Low risk	–	.01
Homko et al (2007) [18]	High risk	+[f]	No
Morrison et al (2001) [24]	/[g]	–	<.001

[a] NICU: neonatal intensive care unit.

[b] CHUMS: Collaborative Home Uterine Monitoring Study.

[c] – indicates less experiences in telemonitoring group versus CG.

[d] NS: not significant.

[e] No exact value is given.

[f] + indicates more experiences in telemonitoring group versus CG.

[g] The slash indicates there was no risk of bias assessment possible for the study because of study design (eg, observational study).

Discussion

The Low Level of Evidence Suggests a Potential Benefit of telemonitoring in Prenatal Care

This review provided a comprehensive description of the use of telemonitoring interventions in obstetrics. Nine of the 14 papers were published from 2007 to present, suggesting that telemonitoring interventions are a relatively new field in obstetrics research. The papers of telemonitoring, which included cervical dilatation or preterm labor as a main outcome, demonstrated that transmitting uterine activity by telecommunication resulted in significant prolonged pregnancy survivals [14,25]. The papers of telemonitoring for GDM demonstrated lower levels of frustration and concerns about their diabetes and a better acceptance of their diabetic condition [17], elated feelings of self-efficacy [18], and a reduction in (unscheduled) face-to-face visits [20,21] in the telemonitoring group compared with the CG. Additionally, a cost reduction [23,24] and elevated feelings of maternal satisfaction [18,26,27] were obtained when telemonitoring was used in obstetrical care. The newborns did have a higher gestational age at delivery [24] and were less likely to be of low birth weight [14,24] or to be admitted to the NICU [14,24] when a telemonitoring group was compared with a CG. Fetuses with abnormal versus normal fetal heart rate at home monitoring were more likely to have an earlier gestational age [16].

Despite the mainly positive results described above, a distinction between studies with low methodological risk of bias assessment and studies with high methodological risk of bias assessment has to be made. When only studies with low risk of bias assessment were taken into account, the added value of telemonitoring became less pronounced. Only pregnant women who transmitted their uterine activity by telecommunication would experience benefits of this technology. They had significant prolonged pregnancy survivals [14,25], and the newborns were less likely to be of low birth weight [14] or to be admitted to the NICU [14]. The study by the Collaborative Home Uterine Monitoring Study (CHUMS) group (1995) was rated low risk for bias but did not mention any significant results for these metrics. On the basis of the low risk for bias criteria, telemonitoring appears to be useful for reducing preterm delivery for pregnant women at risk, but caution should be exercised because only two high-quality studies reported that these benefits were found. Additionally, these papers with a low risk for bias were published in the mid-90s. Their conclusions are questionable when we want to adapt them to current practice because of rapid changes in technology.

Research Gaps and Suggestions for Future Research

Despite the positive results, which are reported above, further research needs to be done to define the added value of telemonitoring and advocate the use of this intervention as a patient management approach in clinical practice. Three main recommendations for future research are made, based on the research gaps elucidated through this review:

1. The level of evidence of the included papers is not high. When a methodological risk of bias is performed, four of these studies classified as *high risk*, one as *medium risk*, and three as *low risk*. Information about randomization (random sequence generation and allocation of concealment) was often lacking; blinding of participants, personnel, and outcomes was not performed in most studies; and none of the used protocols in the intervention groups were available. The level of evidence of the other five studies (which were retrospective studies, a qualitative study, and an observational study) was much lower. There is a need for new multicentric RCTs on different pregnancy conditions in which a blinding for both the patients and the caregivers as the outcomes is performed but with well-considered decisions regarding the ethical aspects. This is to (1) associate the potential of telemonitoring interventions with maternal and fetal outcomes, (2) verify the results which were observed in the mentioned study, (3) investigate the added value of the new technologies nowadays, and (4) improve the evidence on this topic with rigorous research designs.
2. Only four studies reported maternal satisfaction in relation with the use of telemonitoring during their pregnancy (two of them about the use of telemonitoring in pregnancies complicated with GDM and two in the context of labor induction at home). These studies have a relatively small

patient population, ranging from 15 to 70 pregnant women. Patients' satisfaction with the use of telemonitoring systems should be further explored using more robust and validated instruments. Additionally, an evaluation of satisfaction of telemonitoring when used in pregnancies with other pregnancy complications (such as gestational hypertension and premature contractions) and in a bigger patient population is recommended. Alternatively, a thorough qualitative analysis can be conducted to enable an in-depth understanding of patients' satisfaction and the use of that information to improve future technology designs. This may help adjusting the interventions to the target population and can have a positive impact on various domains such as patient compliance and birth experiences.

3. Only two studies performed a cost analysis of prenatal care, including telemonitoring. Both were retrospective studies that were not assessed for risk of bias. Although these studies demonstrated the possibility of cost reduction with the use of telemonitoring, there were visible shortcomings in the study designs. Buysse et al (2008) performed a retrospective study and did not include variables such as time-travel distance from home to hospital and the patient's actual clinical condition. In addition, the staffing costs and equipment costs (based on a reasonable estimate) were not taken into account. Additionally, the data in the study of Morrison et al (2001) were retrospectively collected and did not include the actual clinical condition. In contrast to the previous mentioned study, they asked a fee to finance telemonitoring costs. It is challenging to examine the cost benefit of telemonitoring when it's added to standard prenatal care and whether this is beneficial in both high- and low-risk pregnancies. We recently stated that new technologies can reduce the medicalization of prenatal care [29], but further studies with a prospective design and patient specific treatment(s) are needed to substantiate or reject this hypothesis and to evaluate the cost-effectiveness and health care utilization of telemonitoring in obstetrical care.

Limitations

This review has several limitations that need to be acknowledged. First, the studies were restricted to the English language. Although records written in other languages were excluded, they could be relevant in the scope of this review. Second, a key limitation in the included papers is the heterogeneity of the interventions reported by the investigators. Telemonitoring interventions are frequently multidimensional, containing a range of elements, including the transmission of physiological data, coaching, telephone support, nurse interventions, and Web-based communications [8]. A few studies had a clearly stated aim for the telemonitoring intervention implemented, but, in general, the telemonitoring intervention is poorly described, especially in terms of the assessment of the data transferred and how this assessment leads to a service response or not. Third, the rapid technological advancements that have been seen in the last decade may also impact the ability to compare older and newer studies using different technology. The oldest study dates from 1995, the most recent from 2012. Finally, there was almost no information concerning missing data or the compliance of the patients. The often missing information about compliance rates suggests that telemonitoring regimens may not be appropriate for all patients.

Conclusions

Overall, this review has shown the added value, for both mother and child, of telemonitoring used in a prenatal follow-up program in obstetrical care. However, most of the included studies have a high methodological risk of bias. When only studies with low risk of bias are taken into account, the added value of telemonitoring became less pronounced. Only the pregnant women who transmitted their uterine activity by telecommunication had significant benefits from this technology: they experienced prolonged pregnancy survivals, and the newborns were less likely to be of low birth weight or to be admitted to the NICU. On the basis of the limited results of two high-quality studies conducted in the mid-90s, telemonitoring can be tentatively recommended for pregnant women at risk for preterm delivery. However, more recent RCTs with a blinded protocol and studies investigating patient satisfaction and economic effects in relation to telemonitoring are suggested for future research.

Acknowledgments

This study is part of the Limburg Clinical Research Program (LCRP) UHasselt-ZOL-Jessa, supported by the foundation Limburg Sterk Merk, province of Limburg, Flemish government, Hasselt University, Ziekenhuis Oost-Limburg and Jessa Hospital.

Conflicts of Interest

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this paper. This work was supported by Foundation Mustela (Laureate 2016).

Multimedia Appendix 1

Assessment of risk of bias in included studies.

[PDF File (Adobe PDF File), 84KB - jmir_v19i9e327_app1.pdf]

Multimedia Appendix 2

Summary table of included studies – maternal outcomes.

[PDF File (Adobe PDF File), 41KB - jmir_v19i9e327_app2.pdf]

Multimedia Appendix 3

Summary table of included studies – neonatal outcomes.

[PDF File (Adobe PDF File), 36KB - jmir_v19i9e327_app3.pdf]

References

1. Becker S, Miron-Shatz T, Schumacher N, Krocza J, Diamantidis C, Albrecht UV. mHealth 2.0: experiences, possibilities, and perspectives. JMIR Mhealth Uhealth 2014;2(2):e24 [FREE Full text] [doi: 10.2196/mhealth.3328] [Medline: 25099752]
2. Cruz J, Brooks D, Marques A. Home telemonitoring in COPD: a systematic review of methodologies and patients' adherence. Int J Med Inform 2014 Apr;83(4):249-263. [doi: 10.1016/j.ijmedinf.2014.01.008] [Medline: 24529402]
3. Odibo IN, Wendel PJ, Magann EF. Telemedicine in obstetrics. Clin Obstet Gynecol 2013 Sep;56(3):422-433. [doi: 10.1097/GRF.0b013e318290fef0] [Medline: 23903374]
4. Cowie MR, Lobos AA. Telemonitoring for patients with heart failure. CMAJ 2012 Mar 20;184(5):509-510 [FREE Full text] [doi: 10.1503/cmaj.101456] [Medline: 21464164]
5. Kitsiou S, Paré G, Jaana M. Effects of home telemonitoring interventions on patients with chronic heart failure: an overview of systematic reviews. J Med Internet Res 2015;17(3):e63 [FREE Full text] [doi: 10.2196/jmir.4174] [Medline: 25768664]
6. Giamouzis G, Mastrogiannis D, Koutrakis K, Karayannis G, Parisis C, Rountas C, et al. Telemonitoring in chronic heart failure: a systematic review. Cardiol Res Pract 2012;2012:410820 [FREE Full text] [doi: 10.1155/2012/410820] [Medline: 22720184]
7. Purcell R, McInnes S, Halcomb EJ. Telemonitoring can assist in managing cardiovascular disease in primary care: a systematic review of systematic reviews. BMC Fam Pract 2014 Mar 07;15:43 [FREE Full text] [doi: 10.1186/1471-2296-15-43] [Medline: 24606887]
8. Mushcab H, Kernohan WG, Wallace J, Martin S. Web-based remote monitoring systems for self-managing type 2 diabetes: a systematic review. Diabetes Technol Ther 2015 Jul;17(7):498-509. [doi: 10.1089/dia.2014.0296] [Medline: 25830528]
9. Cruz J, Brooks D, Marques A. Home telemonitoring effectiveness in COPD: a systematic review. Int J Clin Pract 2014 Mar;68(3):369-378. [doi: 10.1111/ijcp.12345] [Medline: 24472009]
10. Gerris J, Delvigne A, Dhont N, Vandekerckhove F, Madoc B, Buyle M, et al. Self-operated endovaginal telemonitoring versus traditional monitoring of ovarian stimulation in assisted reproduction: an RCT. Hum Reprod 2014 Sep;29(9):1941-1948 [FREE Full text] [doi: 10.1093/humrep/deu168] [Medline: 24993931]
11. Pereira I, von Horn K, Depenbusch M, Schultze-Mosgau A, Griesinger G. Self-operated endovaginal telemonitoring: a prospective, clinical validation study. Fertil Steril 2016 Aug;106(2):306-310.e1. [doi: 10.1016/j.fertnstert.2016.03.043] [Medline: 27090865]
12. Gerris J, Geril A, De Sutter P. Patient acceptance of Self-Operated Endovaginal Telemonitoring (SOET): proof of concept. Facts Views Vis Obgyn 2009;1(3):161-170 [FREE Full text] [Medline: 25489461]
13. Higgins JP, Green S, editors. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0. London: The Cochrane Collaboration; 2011.
14. Corwin MJ, Mou SM, Sunderji SG, Gall S, How H, Patel V, et al. Multicenter randomized clinical trial of home uterine activity monitoring: pregnancy outcomes for all women randomized. Am J Obstet Gynecol 1996 Nov;175(5):1281-1285. [Medline: 8942501]
15. The Collaborative Home Uterine Monitoring Study (CHUMS) Group. A multicenter randomized controlled trial of home uterine monitoring: active versus sham device. Am J Obstet Gynecol 1995 Oct;173(4):1120-1127. [Medline: 7485304]
16. Kuleva M, Salomon LJ, Benoist G, Ville Y, Dumez Y. The value of daily fetal heart rate home monitoring in addition to serial ultrasound examinations in pregnancies complicated by fetal gastroschisis. Prenat Diagn 2012 Aug;32(8):789-796. [doi: 10.1002/pd.3903] [Medline: 22653694]
17. Dalfrà MG, Nicolucci A, Lapolla A, TISG. The effect of telemedicine on outcome and quality of life in pregnant women with diabetes. J Telemed Telecare 2009;15(5):238-242. [doi: 10.1258/jtt.2009.081213] [Medline: 19590029]
18. Homko CJ, Santamore WP, Whiteman V, Bower M, Berger P, Geifman-Holtzman O, et al. Use of an internet-based telemedicine system to manage underserved women with gestational diabetes mellitus. Diabetes Technol Ther 2007 Jun;9(3):297-306. [doi: 10.1089/dia.2006.0034] [Medline: 17561800]
19. Homko CJ, Deeb LC, Rohrbacher K, Mulla W, Mastrogiannis D, Gaughan J, et al. Impact of a telemedicine system with automated reminders on outcomes in women with gestational diabetes mellitus. Diabetes Technol Ther 2012 Jul;14(7):624-629 [FREE Full text] [doi: 10.1089/dia.2012.0010] [Medline: 22512287]

20. Pérez-Ferre N, Galindo M, Fernández MD, Velasco V, Runkle I, de la Cruz MJ, et al. The outcomes of gestational diabetes mellitus after a telecare approach are not inferior to traditional outpatient clinic visits. Int J Endocrinol 2010;2010:386941 [FREE Full text] [doi: 10.1155/2010/386941] [Medline: 20628517]
21. Pérez-Ferre N, Galindo M, Fernández MD, Velasco V, de la Cruz MJ, Martín P, et al. A Telemedicine system based on Internet and short message service as a new approach in the follow-up of patients with gestational diabetes. Diabetes Res Clin Pract 2010 Feb;87(2):e15-e17. [doi: 10.1016/j.diabres.2009.12.002] [Medline: 20044162]
22. Brown HL, Britton KA, Brizendine EJ, Hiett AK, Ingram D, Turnquest MA, et al. A randomized comparison of home uterine activity monitoring in the outpatient management of women treated for preterm labor. Am J Obstet Gynecol 1999 Apr;180(4):798-805. [Medline: 10203647]
23. Buysse H, De Moor G, Van Maele G, Baert E, Thienpont G, Temmerman M. Cost-effectiveness of telemonitoring for high-risk pregnant women. Int J Med Inform 2008 Jul;77(7):470-476. [doi: 10.1016/j.ijmedinf.2007.08.009] [Medline: 17923433]
24. Morrison J, Bergauer NK, Jacques D, Coleman SK, Stanziano GJ. Telemedicine: cost-effective management of high-risk pregnancy. Manag Care 2001 Nov;10(11):42-6, 48-9 [FREE Full text] [Medline: 11761593]
25. Wapner RJ, Cotton DB, Artal R, Librizzi RJ, Ross MG. A randomized multicenter trial assessing a home uterine activity monitoring device used in the absence of daily nursing contact. Am J Obstet Gynecol 1995 Mar;172(3):1026-1034. [Medline: 7892843]
26. Rauf Z, O'Brien E, Stampalija T, Ilioniu FP, Lavender T, Alfirevic Z. Home labour induction with retrievable prostaglandin pessary and continuous telemetric trans-abdominal fetal ECG monitoring. PLoS One 2011;6(11):e28129 [FREE Full text] [doi: 10.1371/journal.pone.0028129] [Medline: 22140522]
27. O'Brien E, Rauf Z, Alfirevic Z, Lavender T. Women's experiences of outpatient induction of labour with remote continuous monitoring. Midwifery 2013 Apr;29(4):325-331. [doi: 10.1016/j.midw.2012.01.014] [Medline: 23159160]
28. Baer RJ, Lyell DJ, Norton ME, Currier RJ, Jelliffe-Pawlowski LL. First trimester pregnancy-associated plasma protein-A and birth weight. Eur J Obstet Gynecol Reprod Biol 2016 Mar;198:1-6. [doi: 10.1016/j.ejogrb.2015.12.019] [Medline: 26773241]
29. Gyselaers W, Storms V, Grieten L. New technologies to reduce medicalization of prenatal care: a contradiction with realistic perspectives. Expert Rev Med Devices 2016;13(8):697-699. [doi: 10.1080/17434440.2016.1205484] [Medline: 27336237]

Abbreviations

CG: control group
CHUMS: Collaborative Home Uterine Monitoring Study
COPD: chronic obstructive pulmonary disease
CVD: cardiovascular disease
GDM: gestational diabetes mellitus
NICU: neonatal intensive care unit
NS: not significant
RCT: randomized controlled trial
SD: standard deviation
TM: telemonitoring

Edited by G Eysenbach; submitted 05.01.17; peer-reviewed by D Seale, H Buysse, JB Park, I de la Torre; comments to author 22.03.17; revised version received 01.05.17; accepted 29.07.17; published 27.09.17.

Please cite as:
Lanssens D, Vandenberk T, Thijs IM, Grieten L, Gyselaers W
Effectiveness of Telemonitoring in Obstetrics: Scoping Review
J Med Internet Res 2017;19(9):e327
URL: http://www.jmir.org/2017/9/e327/
doi:10.2196/jmir.7266
PMID:28954715

Original Paper

Barriers and Facilitators to eHealth Use in Daily Practice: Perspectives of Patients and Professionals in Dermatology

Lieneke FM Ariens[1], MD; Florine ML Schussler-Raymakers[1], RN; Cynthia Frima[1], MD; Annebeth Flinterman[2], MD, PhD; Eefje Hamminga[2], MD; Bernd WM Arents[3]; Carla AFM Bruijnzeel-Koomen[1], MD, PhD; Marjolein S de Bruin-Weller[1], MD, PhD; Harmieke van Os-Medendorp[1], PhD

[1]Department of Dermatology and Allergology, University Medical Center Utrecht, Utrecht, Netherlands
[2]Diakonessenhuis, Utrecht, Netherlands
[3]Dutch Association for People with Atopic Dermatitis, Nijkerk, Netherlands

Corresponding Author:
Lieneke FM Ariens, MD
Department of Dermatology and Allergology
University Medical Center Utrecht
Heidelberglaan 100
3584 CX Utrecht
Utrecht, 3584 CX
Netherlands
Phone: 31 887563687
Fax: 31 887555404
Email: l.f.m.ariens@umcutrecht.nl

Abstract

Background: The number of eHealth interventions in the management of chronic diseases such as atopic dermatitis (AD) is growing. Despite promising results, the implementation and use of these interventions is limited.

Objectives: This study aimed to assess opinions of the most important stakeholders influencing the implementation and use of eHealth services in daily dermatology practice.

Methods: The perspectives of health care professionals and patients towards the implementation and use of eHealth services in daily practice were assessed by using a mixed method design. A cross-sectional survey based on the eHealth implementation toolkit (eHit) was conducted to explore factors influencing the adoption of eHealth interventions offering the possibility of e-consultations, Web-based monitoring, and Web-based self-management training among dermatologists and dermatology nurses. The perspectives of patients with atopic dermatitis (AD) regarding the use of eHealth services were discussed in an online focus group.

Results: Health care professionals (n=99) and patients (n=9) acknowledged the value of eHealth services and were willing to use these digital tools in daily dermatology practice. Key identified barriers (statements with <50% of the participants scoring totally agree or agree) in the implementation and adoption of eHealth interventions included concerns about the availability (12/99, 12%) and allocation (14/99, 14%) of resources, financial aspects (26/99, 26%), reliability, security, and confidentially of the intervention itself (29/99, 29%), and the lack of education and training (6/99, 6%).

Conclusions: Health care professionals and patients acknowledge the benefits arising from the implementation and use of eHealth services in daily dermatology practice. However, some important barriers were identified that might be useful in addressing the implementation strategy in order to enhance the implementation success of eHealth interventions in dermatology.

(J Med Internet Res 2017;19(9):e300) doi:10.2196/jmir.7512

KEYWORDS

dermatology; eHealth; implementation; barriers; facilitators

Introduction

In recent years, the increasing prevalence and care of chronic diseases has become a growing burden on modern health care systems. To maintain availability and enhance efficiency of health care, there has been an increasing focus on the development and value of eHealth interventions in the management of chronic diseases [1-5].

Within dermatology, different eHealth interventions are available for use in daily practice. Teledermatology is the most well-known eHealth intervention and has been widely implemented in daily dermatology care [6,7]. The number of health care services combining Web-based accessibility of medical records, systems for interaction between health care professionals and patients (e-consultations), and patient education in the management of chronic skin diseases is growing [8-14]. However, given the promising results on cost-effectiveness, such interventions are not as widely used as might be expected. To implement innovations in daily health care practice, a phase-based approach tailored to specific groups and settings is most successful [15]. The first step of such an approach includes a context analysis to explore factors influencing the implementation. Based on previous research [16-25], different factors affecting the implementation and adoption of eHealth interventions from the perspective of patients and health care providers were identified, including the technological context, product features, and the user and organizational context.

The eHealth implementation toolkit (eHit) is a tool with a phased approach, which can be used to implement eHealth innovations [26]. The design of this tool is based on available evidence about eHealth implementation including data from systematic reviews of barriers and facilitators to implementation of eHealth initiatives, qualitative data derived from interviews of implementers, and the normalization process theory (NPT). The NPT [27] is a conceptual framework for implementing and embedding complex interventions in everyday work. The eHit is described as a sensitizing tool, presenting factors that need consideration before an eHealth intervention can be integrated into daily practice. Assessing attitudes towards the implementation of eHealth interventions and acceptance by relevant stakeholders could provide important information and might enhance the implementation success.

This study aimed to assess the attitudes of relevant stakeholders towards the implementation and adoption of eHealth interventions in daily dermatology practice. A cross-sectional survey based on the eHit [26] was conducted to explore factors affecting the adoption of interventions offering the possibility of e-consultations, Web-based monitoring, and Web-based self-management training among dermatologists and nurses. Patients' participation in their health care is recognized as a key component in the quality of health care. Besides, as an end user of eHealth interventions, patient's involvement at different levels in the implementation process gives valuable insights and may improve the implementation and use of eHealth interventions in daily practice [28,29]. Therefore, perspectives of patients with atopic dermatitis (AD) concerning the use of eHealth interventions were discussed in an online focus group.

Methods

Design

To address perspectives of different stakeholders, a concurrent triangulation mixed method design was used. A cross-sectional survey based on the eHit was performed among Dutch dermatologists and nurses to explore their opinions concerning barriers and facilitators for implementation and adoption of eHealth interventions. In the Netherlands, the following eHealth applications are available for use in daily practice: (1) patient portals offering the possibility of interaction between health care professionals and patients (e-consultations); (2) Web-based monitoring of the disease by using digital photographs and insight in the medical record; and (3) self-management trainings focusing on treatment adherence, prevention of exacerbations, and coping with itch and psychological problems. Furthermore, perspectives of adult AD patients and parents of children with AD regarding the use of such eHealth interventions in daily dermatology practice were discussed in a qualitative study in an online focus group.

Sampling and Recruitment

Dutch dermatologists were purposively recruited to participate in the survey anonymously, via an email invitation by the authors sent by the Dutch Association for Dermatologists (NVDV) or personal email invitation. Members of a platform for nurses, which is aimed to increase the expertise of nursing care for patients with AD or itch, were approached to be included in this study. Additionally, dermatologists and nurses were recruited from an academic and regional hospital in the Netherlands, participating in an implementation project for digital care in patients with AD.

The qualitative part of this study contained an online focus group including adult AD patients and parents of young children with AD (aged 0-8 years). Participants were recruited by 2 research nurses at the outpatient clinic of the dermatology department of the University Medical Center Utrecht (UMCU) and through an advertisement by the Dutch Association for People with Atopic Dermatitis (VMCE). Purposive sampling was performed to include participants with and without experience with digital tools.

Outcome Parameters

EHit Survey

A survey based on the eHit [26] was developed to explore barriers and facilitators for implementation and adoption of an eHealth intervention among dermatologists and nurses. This Web-based questionnaire contained 23 statements with a 5-point Likert scale ranging from 1 (completely disagree) to 5 (completely agree), which were grouped into 3 main sections.

1. Context: national and hospital policy, attitudes of professionals, resources, and risks
2. Intervention: consequences for professional – patient relationship, safety, ease of use, and benefits and cost–effectiveness

3. Workforce: consequences for work load, collaboration, work patterns, and training and responsibility.

The survey used in this study based on the eHit was not validated.

Topics and Procedure in the Online Focus Group

Perspectives of AD patients and parents of children with AD were explored by using an online focus group. In 2 weeks, 8 statements were posted in the online focus group on Facebook to explore patients' opinions concerning experiences with digital tools in health care, the usefulness of digital contact with other patients and e-learning, advantages and disadvantages of digital tools, and the willingness and requirements for the use of such digital tools in daily practice.

The online focus group was set up in a closed account on Facebook for a 2-week period. After the informed consent was obtained, participants were invited to join the online focus group by the researcher or research nurse. The discussion in the online focus group was started by the researcher giving an explanation considering the aim of the discussion and posting the first statement. Participants were motivated to react on the statements and to participate actively in the group discussion. The researcher and research nurse asked for clarifications and experiences, asked questions, and made summaries. The researcher and research nurse were both involved in the development and research of eHealth interventions. The research nurse was also involved in the patient care of two participants however the researcher was not connected with the participants. The quantitative and qualitative part of this study did not fall under the scope of the Medical Research Involving Human Subjects, which was confirmed by the Medical Research Ethics Committee of the UMC Utrecht, the Netherlands for the qualitative part including patients.

Analysis

Results of the eHit survey were analyzed using IBM SPSS Statistics Version 21 (SPSS Inc, Chicago, IL, USA). Answers on the 5-point Likert scale ranging from 1 (completely disagree) to 5 (completely agree) were categorized into 3 categories: totally disagree or disagree, totally agree or agree, and unknown. The mean number of participants who totally agreed or agreed was calculated per statement of each category (context, intervention, and workforce). Statements with <50% of the participants scoring totally agree or agree were considered to be a barrier for implementation and adoption of an eHealth intervention. Statements with >50% of the participants scoring totally agree or agree were considered to be a facilitator. Results are shown for the total group of professionals and categorized in professionals with and without eHealth experience. Differences in responses between health care professionals with and without eHealth experience, medical doctors, and nurses were analyzed by using the Pearson's chi-squared test.

A generic qualitative analysis was used to analyze the discussion yield in the online focus group including coding, categorizing, formulizing themes, and connecting and interpreting them [30]. To ensure trustworthiness, interpretations and conclusions were summarized before closing the online focus group to conduct a member check. The researcher and research nurse both coded all text and discussed the analyses and results until consensus was achieved.

Results

EHit Survey

In total 800 health care professionals were approached to complete the survey, of which 108 responded, yielding a response rate of 14%. The survey response rates among members of de NVDV, the nurses' platform, and the participating hospitals were 48/670 (7%), 24/69 (35%), and 36/61 (59%), respectively. Reasons for not responding on the questionnaire are unknown. Respondents with another profession than dermatologist or nurse were excluded from the analysis. As shown in Table 1, 99 health care professionals who completed the eHit survey were included of which 65 (66%) were dermatologists and 34 nurses (34%). Out of the 99 participants, 65 (66%) were female and the mean age (SD) of the total group of professionals was 47 (SD 10.5). A total of 26 dermatologists and 16 nurses reported no experience in digital care compared to 39 dermatologists and 18 nurses with experience in digital care. The 23 statements used in the survey and the respondents scores on agreement are fully presented in Multimedia Appendix 1.

The mean number of participants that totally agreed with statements related to the context was 36% (Figure 1). Context related barriers (<50% of the respondents scoring totally agree or agree) included the availability of resources (12 /99, 12%), their allocation (14/99, 14%), security policy (31/99, 31%), and the organizational culture welcoming eHealth initiatives (33/99, 33%). Context related facilitators (≥50% of the respondents scoring totally agree or agree) included "fits in local policy regarding efficiency and patient-centered care" (53/99, 54%) and "good working relationships between different professionals involved in the implementation of digital care" (51/99, 52%). Among the professionals with eHealth experience, the highest scoring facilitator was "presence of particular opinion leaders who are likely to support the implementation of digital care" (36/57, 63%).

A total of 47% of respondents totally agreed with statements related to the intervention (Figure 2). Credibility of digital care in terms of confidentiality, security, and reliability of the intervention (29/99, 29%), ease of use for patients (35/99, 35%) and professionals (37/99, 37%), and the cost-effectiveness of eHealth interventions (26/99, 26%) were identified as barriers for implementation. Facilitators included "facilitation of health care professional – patient interaction" (59/99, 60%), benefits for patients (61/99, 62%), and the notion of professionals that they do not fear losing control when using e-consultations instead of face-to-face visits (70/99, 71%).

Table 1. Characteristics of respondents.

Characteristics	Total (n=99)	No experience with digital care (n=42)		Experience with digital care (n=57)	
		Dermatologist (n=26)	Nurse (n=16)	Dermatologist (n=39)	Nurse (n=18)
Gender, n (%)					
Male	34 (34)	12 (46)	0 (0)	22 (56)	0 (0)
Female	65 (66)	14 (54)	16 (100)	17 (44)	18 (100)
Age (missing n=1), mean (SD)	47.0 (10.5)	45.0 (12.6)	50.4 (6.6)	46.1 (10.9)	48.7 (8.8)
Organization, n (%)					
General hospital	50 (51)	12 (46)	12 (75)	16 (41)	10 (56)
University hospital	37 (37)	13 (50)	2 (13)	16 (41)	6 (33)
Primary care or home health care or other	5 (5)	0 (0)	2 (13)	1 (3)	2 (11)
Independent treatment centre	7 (7)	1 (4)	0 (0)	6 (15)	0 (0)
Experience with digital care, n (%)					
E-consultation	23 (23)	0 (0)	0 (0)	15 (39)	8 (44)
Patient portal	16 (16)	0 (0)	0 (0)	7 (18)	9 (50)
Web-based self-management training	14 (14)	0 (0)	0 (0)	4 (10)	10 (56)
Teledermatology	18 (18)	0 (0)	0 (0)	18 (46)	0 (0)
Video of webcam consult	2 (2)	0 (0)	0 (0)	1 (3)	1 (6)
Other[a]	9 (9)	0 (0)	0 (0)	8 (21)	1 (6)

[a] Other eHealth tools such as wound-telemonitoring system, portal for Web-based questionnaires, informational site.

Figure 1. Percentage of participants scoring totally agree or agree on statements related to the context.

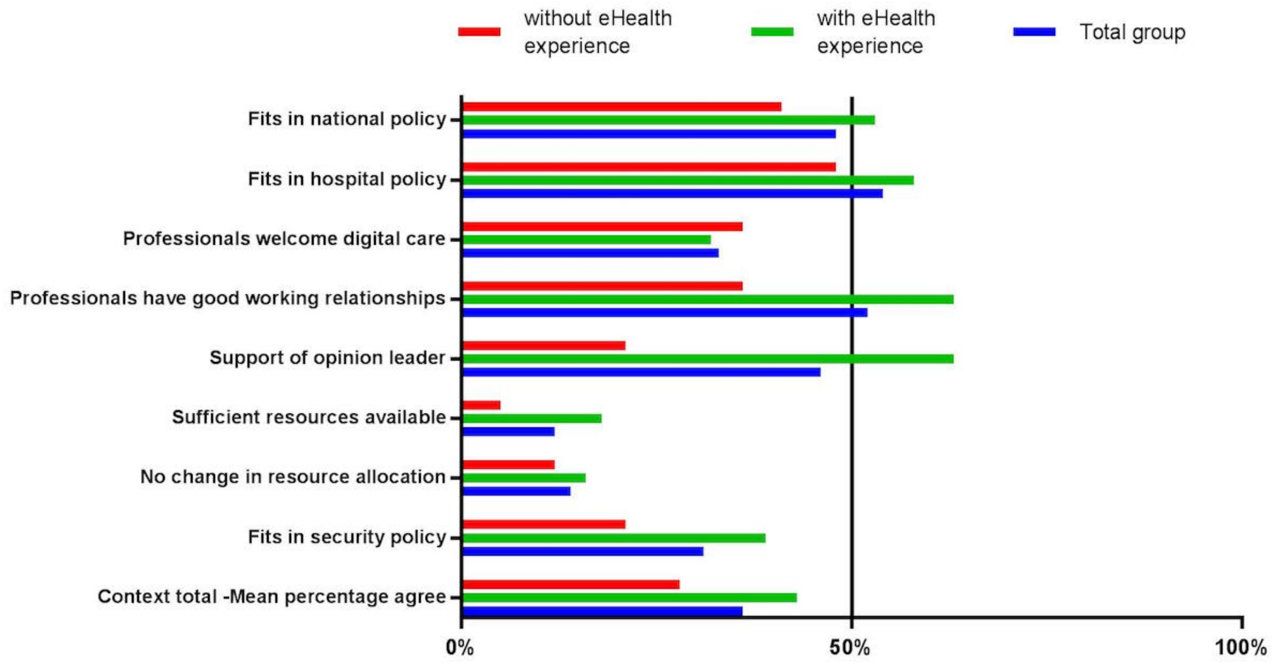


Figure 2. Percentage of participants scoring totally agree or agree on statements related to the intervention.

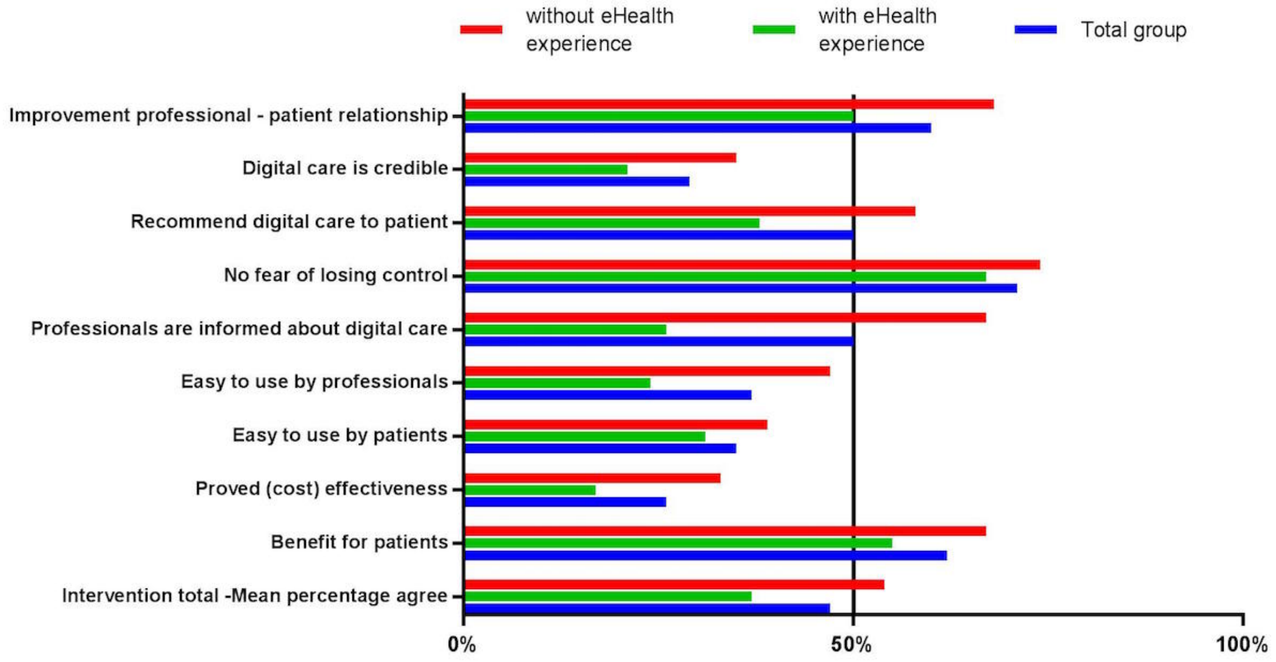
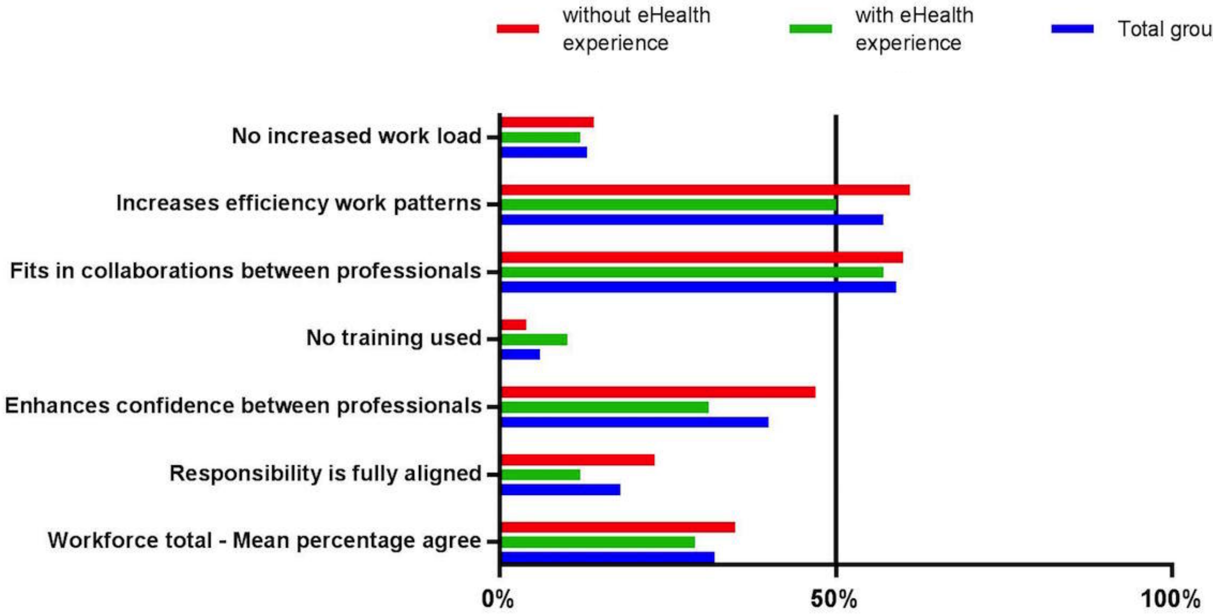


Figure 3. Percentage of participants scoring totally agree or agree on statements related to the workforce.



Finally, the percentage of participants totally agreeing with statements related to workforce comprised 32% (Figure 3). Barriers for implementation related to workforce included concerns about increased workload (13/99, 13%), the need for training prior to the implementation of digital care (6/99, 6%), and the alignment of responsibility for the use of digital care (18/99, 18%). The possible increase of efficiency in work patterns (56/99, 57%) and collaboration between health care professionals (58/99, 59%) were identified as facilitators.

As shown in Multimedia Appendix 1, the respondents' scores on agreement significantly differed between professionals with and without experience in using eHealth interventions for the statements "nurses and medical doctors involved in the implementation of digital care have good working relationships, good communication and co-operation" (*P*=.01), and "there are particular opinion leaders within the hospital who are likely to support the implementation of digital care" (*P*<.001) related to the context. Agreement scores related to the intervention were significantly different among experienced and non-experienced professionals for the statements "I will recommend my patients to use digital care" (*P*=.004), "I know what digital care comprises and how it can be used" (*P*<.001), "digital care is easy to use by medical doctors and nurses" (*P*=.001) and "digital care has been well evaluated and has been demonstrated to improve health care in a cost effective manner" (*P*=.02). Considering workforce, the agreement of health care professionals on the statement "workload during (future) implementation is not increased" (*P*=.008) significantly differed between experienced and nonexperienced professionals. Professionals with experience in digital care were more likely to totally agree with the above-mentioned statements compared to professionals without experience in digital care.

Online Focus Group

A total of 15 AD patients or parents of children with AD were approached to participate in the online focus group. No response to the invitation, lack of experience with Facebook, and privacy concerns were the main reasons not to participate in the online focus group. In total, 7 patients with AD and 2 parents of children with AD were included (Table 2). Out of the 9 participants, 8 were female and 1 was male.

Digital Tools in Health Care: Advantages and Disadvantages

Most patients were not experienced in using digital tools. However, discussing the value of e-consultation as a potential digital tool, all patients acknowledged its value and were willing to use these tools in daily practice. Reported advantages included the possibility of sending photographs and request e-repeat prescriptions, the quick and satisfying response to questions, and the fact that it is available on any weekday without making an appointment. The possibility of contacting a health care professional by using e-consulting at the time of exacerbations is the most important advantage reported by the participants. By using e-consultations, patients expect to receive effective care at the moment it is most useful and needed. Sending photographs makes it possible for health care professionals to access disease activity and give treatment advice. Furthermore, an e-consultation can eliminate an extra face-to-face contact. It is timesaving and patient friendly:

> *...e-consultations are a valuable addition. It is not to cut costs but to increase the quality of treatment.* [ID04]

Insight in the medical record as one of the functionalities in the electronic patients' portal was appreciated by patients who used this digital tool. Patients acknowledged the possibility to review the altered treatment again.

> *The advantage is that you can read the report again if needed.* [ID06]

The possibility to gain insight in their medical record and read the status reported by the caregiver made them feel more confident with their disease and treatment.

Participants also reported disadvantages of the Web-based tools. Patients think that there will be less personal contact with the caregiver:

> *A disadvantage of using an e- consultation is that you will increasingly depend on a computer to control your disease via the Internet. The personal (social) contact with your caregiver decreases.* [ID07]

Besides, patients are not willing to pay extra costs for this digital tool and have concerns about the privacy and safety of digital care. They see the possibility of using an e-consultation as complementary, not as a replacement of face-to-face consultations: "I think a healthy alternation between face-to-face consultations and e- consultations is necessary." (ID09)

Other Digital Tools: Digital Contact With Other Patients and E-Learning

Participants acknowledged the value of having digital contact with other patients by using a digital forum or chat area. They mentioned the usefulness of sharing experiences with fellow patients, especially in combination with the medical point of view from the caregiver. Patients endorsed the value of the Web-based e-learning program substance but noticed there is already a lot of information available on the Internet. Websites referred by the caregiver were considered to be more reliable:

> *If a caregiver refers to a website, it's more reliable in my eyes.* [ID05]

Table 2. Characteristics of participants of the online focus group.

Characteristics	ID	Sex	Age	Educational level[a]	Internet use	Internet skills
University hospital - Children's department						
	1	F	37	Medium	(almost) daily	Good
	2	F	33	High	(almost) daily	Good
University hospital - Adult's department						
	3	F	19	Medium	(almost) daily	Good
	4	F	24	High	(almost) daily	Moderate
	5	F	55	Low	Several times a week	Moderate
Dutch Association for People with Atopic Dermatitis (VMCE)						
	6	F	34	High	(almost) daily	Very good
	7	F	45	Low	(almost) daily	Good
	8	F	21	Medium	(almost) daily	Good
	9	M	59	High	(almost) daily	Very good

[a] Educational level: low, elementary education, high school or middle-level applied education; high: higher professional or academic education.

Discussion

The results of this study illustrate the overall positive attitude of health care professionals and patients towards the implementation and use of eHealth services offering the possibility of e-consultations, Web-based monitoring, and Web-based self-management training in daily dermatology practice. However, some remaining challenges were also identified. Both health care professionals and patients acknowledge the value of eHealth services and are willing to use such interventions in daily practice. Patients appreciate the comprehensive accessibility of digital care and the possibility to gain insight in the medical record and to contact a health care professional in times of exacerbations. Besides, health care professionals value the potential increase of efficiency in work patterns and collaborations. Key identified barriers in the implementation and adoption of eHealth interventions included concerns about financial aspects, reliability, security, confidentiality, and cost-effectiveness of the intervention itself as well as the lack of education and training.

These barriers are in line with findings from current literature [18-21,23,31-33]. In different studies investigating the implementation and adoption of a variety of eHealth interventions in a range of settings, technological knowledge, and skills: financial aspects, social and organizational support, and the lack of education and training are the most frequently noted barriers [18-21,23,31-33]. Crucial factors leading to the successful implementation of teledermatology have been the focus on embedding the intervention in the existing health care infrastructure, the comprehensive support and training of future users, and the clear assignment of persons who took full responsibility for the entire process [34]. Besides, user satisfaction was identified as an important factor in the implementation process [34]. These findings demonstrate that, despite methodological differences, studies identified quite similar factors that should be considered before implementing eHealth interventions in daily practice.

An interesting finding of this study is the more positive attitude towards implementation and use of eHealth services among professionals with experience in using eHealth applications. Experienced health care professionals acknowledge the advantages arising from eHealth services and report fewer barriers in the implementation process as compared to those who are not experienced in using such services. Previous research investigating the opinions of health care professional towards the introduction of a new eHealth service in Sweden also showed a significantly more positive attitude among experienced professionals compared to inexperienced professionals [35]. These findings imply that professionals need to overcome some obstacles to enable them to recognize potential benefits that can be derived from implementation of eHealth interventions. Receiving education and training in the implementation process might help to lower the threshold and increase acceptance of eHealth interventions in daily dermatology practice [15].

Results of the online focus group demonstrate a positive attitude towards to adoption and use of eHealth interventions in daily dermatology practice among AD patients. Previous research demonstrates that the acceptance of eHealth interventions among patients suffering from chronic diseases relies on their attitude towards the usefulness of Internet in personal health care [24]. In this study, most patients were not experienced in using eHealth applications. Interestingly, a lack of experience seemed not to negatively influence the acceptance of eHealth interventions. Compared with other studies assessing the attitude of patients towards the introduction of eHealth services [21-25,33], we found a relatively positive attitude towards the use of digital use in daily practice. This might be explained by the recruitment of a relatively small group of patients in a tertiary hospital and through an advertisement by the Dutch Association for People with Atopic Dermatitis (VMCE). It is possible that patients who are interested in eHealth were more likely to respond, resulting in an overestimation of the positive attitude. Besides, one of the research nurses recruiting patients for the online focus group at the outpatient clinic of the dermatology department of UMCU was also involved in the patient care of 2 participants. Therefore, potential sample selection bias might have been introduced in this group. Additionally, patients recruited in a tertiary hospital are likely to have a more severe AD, which might have influenced their attitude towards the adoption and use of eHealth services.

In this study, the eHit was used to develop a survey to explore the attitude of health care professionals. The eHit is a tool with a phased approach that was designed based on present evidence about the implementation of eHealth interventions [26]. The use of the eHit was considered to be feasible and acceptable by a variety of professionals and use in different health care systems [36]. Therefore, the results of this survey combined with the results from the online focus group provide a good reflection of factors that need consideration before implementing and embedding interventions in daily practice.

A research limitation of the present study is the low response rate to the eHit survey, which may not fully represent the target population. Besides, the highest response rate was found among professionals recruited from hospitals participating in an implementation project for digital care in patients with AD. Possibly, health care professionals who are interested in eHealth or are already working with such interventions might be more likely to complete the survey, leading to response bias.

In conclusion, this study attempts to use a systematical method to provide attitudes and factors influencing the implementation and adoption of eHealth services in daily dermatology practice of key stakeholders. The overall attitude towards the adoption and use of eHealth services among health care professionals and AD patients was positive; however, we also identified some important challenges in the implementation process. Findings of this study might be useful in addressing the implementation strategy to the health care professionals' and patients' values, needs, and attitudes and consequently enhance the implementation success of eHealth interventions in daily dermatology.

Acknowledgments

The authors thank the nurses of the platform of nursing care for patients with AD or itch for the help with data collection. The authors also thank the Dutch Association for Dermatologists (NVDV) for the help with data collection. This work was supported by ZonMw grant number 945062181.

Authors' Contributions

LFMA, FMLF, CF, and HOM contributed to conception and study design, data analysis, interpretation of data and drafting the article. AEF, EAH, BWMA, CAFMB-K, and MSB-W contributed to acquisition of data, interpretation of data and revising the article critically for important intellectual content. Each author listed on the manuscript has seen and approved the submission of this version of the manuscript and takes full responsibility for the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Results of the eHit survey.

[PDF File (Adobe PDF File), 71KB - jmir_v19i9e300_app1.pdf]

References

1. Hutchesson MJ, Rollo ME, Krukowski R, Ells L, Harvey J, Morgan PJ, et al. eHealth interventions for the prevention and treatment of overweight and obesity in adults: a systematic review with meta-analysis. Obes Rev 2015 May;16(5):376-392. [doi: 10.1111/obr.12268] [Medline: 25753009]
2. Zhao J, Zhai Y, Zhu W, Sun D. Effectiveness of telemedicine for controlling asthma symptoms: a systematic review and meta-analysis. Telemed J E Health 2015 Jun;21(6):484-492. [doi: 10.1089/tmj.2014.0119] [Medline: 25393915]
3. Kotb A, Cameron C, Hsieh S, Wells G. Comparative effectiveness of different forms of telemedicine for individuals with heart failure (HF): a systematic review and network meta-analysis. PLoS One 2015;10(2):e0118681 [FREE Full text] [doi: 10.1371/journal.pone.0118681] [Medline: 25714962]
4. Elbert NJ, van Os-Medendorp O, van Renselaar W, Ekeland AG, Hakkaart-van Roijen L, Raat H, et al. Effectiveness and cost-effectiveness of ehealth interventions in somatic diseases: a systematic review of systematic reviews and meta-analyses. J Med Internet Res 2014;16(4):e110 [FREE Full text] [doi: 10.2196/jmir.2790] [Medline: 247739471]
5. Black AD, Car J, Pagliari C, Anandan C, Cresswell K, Bokun T, et al. The impact of eHealth on the quality and safety of health care: a systematic overview. PLoS Med 2011 Jan;8(1):e1000387 [FREE Full text] [doi: 10.1371/journal.pmed.1000387] [Medline: 21267058]
6. Coates SJ, Kvedar J, Granstein RD. Teledermatology: from historical perspective to emerging techniques of the modern era: part I: history, rationale, and current practice. J Am Acad Dermatol 2015 Apr;72(4):563-74; quiz 575. [doi: 10.1016/j.jaad.2014.07.061] [Medline: 25773407]
7. Coates SJ, Kvedar J, Granstein RD. Teledermatology: from historical perspective to emerging techniques of the modern era: part II: emerging technologies in teledermatology, limitations and future directions. J Am Acad Dermatol 2015 Apr;72(4):577-86; quiz 587. [doi: 10.1016/j.jaad.2014.08.014] [Medline: 25773408]
8. Chambers CJ, Parsi KK, Schupp C, Armstrong AW. Patient-centered online management of psoriasis: a randomized controlled equivalency trial. J Am Acad Dermatol 2012 Jun;66(6):948-953 [FREE Full text] [doi: 10.1016/j.jaad.2011.05.047] [Medline: 21890236]
9. Armstrong AW, Johnson MA, Lin S, Maverakis E, Fazel N, Liu F. Patient-centered, direct-access online care for management of atopic dermatitis: a randomized clinical trial. JAMA Dermatol 2015 Feb;151(2):154-160. [doi: 10.1001/jamadermatol.2014.2299] [Medline: 25338198]
10. Santer M, Muller I, Yardley L, Burgess H, Selinger H, Stuart BL, et al. Supporting self-care for families of children with eczema with a Web-based intervention plus health care professional support: pilot randomized controlled trial. J Med Internet Res 2014 Mar 04;16(3):e70 [FREE Full text] [doi: 10.2196/jmir.3035] [Medline: 24594972]
11. van Os-Medendorp H, Koffijberg H, Eland-de Kok PC, van der Zalm A, de Bruin-Weller MS, Pasmans SG, et al. E-health in caring for patients with atopic dermatitis: a randomized controlled cost-effectiveness study of internet-guided monitoring and online self-management training. Br J Dermatol 2012 May;166(5):1060-1068. [doi: 10.1111/j.1365-2133.2012.10829.x] [Medline: 22268960]
12. de Graaf M, Totte J, Breugem C, van Os-Medendorp H, Pasmans S. Evaluation of the compliance, acceptance, and usability of a web-based ehealth intervention for parents of children with infantile hemangiomas: usability study. JMIR Res Protoc 2013 Dec 17;2(2):e54 [FREE Full text] [doi: 10.2196/resprot.2897] [Medline: 24345450]

13. van Os-Medendorp H, van Leent-de Wit I, de Bruin-Weller M, Knulst A. Usage and users of online self-management programs for adult patients with atopic dermatitis and food allergy: an explorative study. JMIR Res Protoc 2015 May 23;4(2):e57 [FREE Full text] [doi: 10.2196/resprot.4134] [Medline: 25998229]
14. Bergmo TS, Wangberg SC, Schopf TR, Solvoll T. Web-based consultations for parents of children with atopic dermatitis: results of a randomized controlled trial. Acta Paediatr 2009 Feb;98(2):316-320. [doi: 10.1111/j.1651-2227.2008.01033.x] [Medline: 18795905]
15. Grol R, Grimshaw J. From best evidence to best practice: effective implementation of change in patients' care. Lancet 2003 Oct 11;362(9391):1225-1230. [doi: 10.1016/S0140-6736(03)14546-1] [Medline: 14568747]
16. Boddy D, King G, Clark JS, Heaney D, Mair F. The influence of context and process when implementing e-health. BMC Med Inform Decis Mak 2009 Jan 30;9:9 [FREE Full text] [doi: 10.1186/1472-6947-9-9] [Medline: 19183479]
17. Joseph V, West RM, Shickle D, Keen J, Clamp S. Key challenges in the development and implementation of telehealth projects. J Telemed Telecare 2011;17(2):71-77. [doi: 10.1258/jtt.2010.100315] [Medline: 21097563]
18. May CR, Finch TL, Cornford J, Exley C, Gately C, Kirk S, et al. Integrating telecare for chronic disease management in the community: what needs to be done? BMC Health Serv Res 2011;11:131 [FREE Full text] [doi: 10.1186/1472-6963-11-131] [Medline: 21619596]
19. Postema TR, Peeters JM, Friele RD. Key factors influencing the implementation success of a home telecare application. Int J Med Inform 2012 Jun;81(6):415-423 [FREE Full text] [doi: 10.1016/j.ijmedinf.2011.12.003] [Medline: 22226925]
20. Liu C. Key factors influencing the intention of telecare adoption: an institutional perspective. Telemed J E Health 2011 May;17(4):288-293. [doi: 10.1089/tmj.2010.0184] [Medline: 21480787]
21. Brooks E, Manson SM, Bair B, Dailey N, Shore JH. The diffusion of telehealth in rural American Indian communities: a retrospective survey of key stakeholders. Telemed J E Health 2012;18(1):60-66 [FREE Full text] [doi: 10.1089/tmj.2011.0076] [Medline: 22082106]
22. Sanders C, Rogers A, Bowen R, Bower P, Hirani S, Cartwright M, et al. Exploring barriers to participation and adoption of telehealth and telecare within the whole system demonstrator trial: a qualitative study. BMC Health Serv Res 2012;12:220 [FREE Full text] [doi: 10.1186/1472-6963-12-220] [Medline: 22834978]
23. Rogove HJ, McArthur D, Demaerschalk BM, Vespa PM. Barriers to telemedicine: survey of current users in acute care units. Telemed J E Health 2012;18(1):48-53. [doi: 10.1089/tmj.2011.0071] [Medline: 22082107]
24. Duplaga M. A cross-sectional study assessing determinants of the attitude to the introduction of eHealth services among patients suffering from chronic conditions. BMC Med Inform Decis Mak 2015 Apr 19;15:33 [FREE Full text] [doi: 10.1186/s12911-015-0157-3] [Medline: 25927312]
25. Sun N, Rau PP. The acceptance of personal health devices among patients with chronic conditions. Int J Med Inform 2015 Apr;84(4):288-297. [doi: 10.1016/j.ijmedinf.2015.01.002] [Medline: 25655783]
26. Murray E, May C, Mair F. Development and formative evaluation of the e-health implementation toolkit (e-HIT). BMC Med Inform Decis Mak 2010;10:61 [FREE Full text] [doi: 10.1186/1472-6947-10-61] [Medline: 20955594]
27. 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;4:29 [FREE Full text] [doi: 10.1186/1748-5908-4-29] [Medline: 19460163]
28. van Bruinessen IR, van Weel-Baumgarten EM, Snippe HW, Gouw H, Zijlstra JM, van Dulmen S. Active patient participation in the development of an online intervention. JMIR Res Protoc 2014;3(4):e59 [FREE Full text] [doi: 10.2196/resprot.3695] [Medline: 25379679]
29. Nordin C, Michaelson P, Eriksson MK, Gard G. It's about me: patients' experiences of patient participation in the web behavior change program for activity in combination with multimodal pain rehabilitation. J Med Internet Res 2017 Jan 18;19(1):e22 [FREE Full text] [doi: 10.2196/jmir.5970] [Medline: 28100440]
30. Holloway I, Wheeler S. Qualitative Research in Nursing. Hoboken, NJ: Wiley; 2002.
31. Uscher-Pines L, Kahn JM. Barriers and facilitators to pediatric emergency telemedicine in the United States. Telemed J E Health 2014 Nov;20(11):990-996 [FREE Full text] [doi: 10.1089/tmj.2014.0015] [Medline: 25238565]
32. Duplaga M. Searching for a role of nursing personnel in developing landscape of ehealth: factors determining attitudes toward key patient empowering applications. PLoS One 2016;11(4):e0153173 [FREE Full text] [doi: 10.1371/journal.pone.0153173] [Medline: 27049525]
33. Valenzuela EA, De Smedt A, Guldolf K, Vandervorst F, Van Hooff R, Fernandez TH, et al. Opinions and beliefs about telemedicine for emergency treatment during ambulance transportation and for chronic care at home. Interact J Med Res 2016 Mar 30;5(1):e9 [FREE Full text] [doi: 10.2196/ijmr.5015] [Medline: 27029999]
34. Tensen E, van der Heijden JP, Jaspers MW, Witkamp L. Two decades of teledermatology: current status and integration in national healthcare systems. Curr Dermatol Rep 2016;5:96-104 [FREE Full text] [doi: 10.1007/s13671-016-0136-7] [Medline: 27182461]
35. Ålander T, Scandurra I. Experiences of healthcare professionals to the introduction in Sweden of a public ehealth service: patients' online access to their electronic health records. Stud Health Technol Inform 2015;216:153-157. [Medline: 26262029]

36. MacFarlane A, Clerkin P, Murray E, Heaney DJ, Wakeling M, Pesola U, et al. The e-health implementation toolkit: qualitative evaluation across four European countries. Implement Sci 2011 Nov 19;6:122 [FREE Full text] [doi: 10.1186/1748-5908-6-122] [Medline: 22098945]

Abbreviations

AD: atopic dermatitis
eHit: eHealth implementation toolkit
NVDV: Dutch Association for Dermatologists
UMCU: University Medical Center Utrecht
VMCE: Dutch Association for People with Atopic Dermatitis (VMCE)

Edited by CL Parra-Calderón; submitted 15.02.17; peer-reviewed by P Soyer, T Schopf, L Witkamp, M Duplaga, D Nault; comments to author 30.03.17; revised version received 24.05.17; accepted 18.06.17; published 05.09.17.

Please cite as:
Ariens LFM, Schussler-Raymakers FML, Frima C, Flinterman A, Hamminga E, Arents BWM, Bruijnzeel-Koomen CAFM, de Bruin-Weller MS, van Os-Medendorp H
Barriers and Facilitators to eHealth Use in Daily Practice: Perspectives of Patients and Professionals in Dermatology
J Med Internet Res 2017;19(9):e300
URL: http://www.jmir.org/2017/9/e300/
doi:10.2196/jmir.7512
PMID:28874336

Original Paper

The Digital Divide and Health Disparities in China: Evidence From a National Survey and Policy Implications

Y Alicia Hong[1,2*], PhD; Zi Zhou[1*], PhD; Ya Fang[1], MD, PhD; Leiyu Shi[3], DrPH

[1]School of Public Health, Xiamen University, Xiamen, China
[2]School of Public Health, Texas A&M University, College Station, TX, United States
[3]Department of Health Policy and Management, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD, United States
[*]these authors contributed equally

Corresponding Author:
Y Alicia Hong, PhD
School of Public Health
Xiamen University
No. 4221-117, Xiang'an South Road
Xiamen, 361102
China
Phone: 86 9794369343
Fax: 86 9794584267
Email: yhong@sph.tamhsc.edu

Abstract

Background: The digital divide persists despite broad accessibility of mobile tools. The relationship between the digital divide and health disparities reflects social status in terms of access to resources and health outcomes; however, data on this relationship are limited from developing countries such as China.

Objective: The aim of this study was to examine the current rates of access to mobile tools (Internet use and mobile phone ownership) among older Chinese individuals (aged ≥45 years), the predictors of access at individual and community levels, and the relationship between access to mobile tools and health outcomes.

Methods: We drew cross-sectional data from a national representative survey, the China Health and Retirement Longitudinal Study (CHARLS), which focused on the older population (aged ≥45 years). We used two-level mixed logistic regression models, controlling for unobserved heterogeneity at the community and individual levels for data analysis. In addition to individual-level socioeconomic status (SES), we included community-level resources such as neighborhood amenities, health care facilities, and community organizations. Health outcomes were measured by self-reported health and absence of disability based on validated scales.

Results: Among the 18,215 participants, 6.51% had used the Internet in the past month, and 83% owned a mobile phone. In the multivariate models, Internet use was strongly associated with SES, rural or urban residence, neighborhood amenities, community resources, and geographic region. Mobile phone ownership was strongly associated with SES and rural/urban residence but not so much with neighborhood amenities and community resources. Internet use was a significant predictor of self-reported health status, and mobile phone ownership was significantly associated with having disability even after controlling for potential confounders at the individual and community levels.

Conclusions: This study is one of the first to examine digital divide and its relationship with health disparities in China. The data showed a significant digital divide in China, especially in the older population. Internet access is still limited to people with higher SES; however, the mobile phone has been adopted by the general population. The digital divide is associated with not only individual SES but also community resources. Future electronic health (eHealth) programs need to consider the accessibility of mobile tools and develop culturally appropriate programs for various social groups.

(J Med Internet Res 2017;19(9):e317) doi:10.2196/jmir.7786

KEYWORDS

digital divide; health disparities; Internet; mobile phone; China

Introduction

Over the past decade, global access to mobile technologies such as the Internet and cell phones has dramatically increased. Such access has transformed the way people receive information and communicate with one another; in fact, prior studies have shown that those with access to mobile technologies have better mental health, physical health, and medical decision-making skills [1,2]. Accompanying the "Internet of Things" and the "e-lifestyle" is the persistent digital divide, defined as the gap between those with access to new forms of information technology and those without. [3-5]. Even in the Western countries with high penetration rates of mobile tools, digital inequality remains significant [6-8], and the digital divide strongly correlates with health disparities [9,10].

Literature on the relationship of the digital divide and health disparities suggests that an individual's lifestyle choices, including the use of mobile technologies, are not autonomous but are instead constrained or enabled by one's social status and access to economic and other resources [3,4,11-13]. Having access to resources including mobile technologies enables an individual, family, or social group to receive more up-to-date health information, to obtain social support, to adopt healthy behaviors, and to make more informed medical decisions, and therefore, have better health outcomes; this in turn reshapes their socioeconomic status (SES; [3,10,13]). Empirical data corroborated this theory and showed that both the digital divide and health disparities are predicted by older age, low level of education, low income, racial or ethnic minority status, and rural residence [3-13].

Current literature on digital divide and health disparities is limited in three aspects. First, existing studies on predictors of the digital divide and its relationship with health disparities have been limited to individual-level SES factors such as age, race/ethnicity, gender, education, and income; however, community-level factors such as neighborhood characteristics and community resources have rarely been examined. Sociologists have long documented the impact of community on individual health and well-being [14-16]. Second, only a few studies on the digital divide and health disparity have been conducted in older adults, who are often the last group to adopt technology and more likely to face health disparities [8,17,18]. In many countries, the elder population is the fastest growing population but has been left out of the accelerative "e-lifestyle" movement [19-22]. Third, most existing studies on the digital divide and health disparities were conducted in Western countries, with limited data from developing countries, despite the rapid adoption rates of mobile technologies in these countries [23,24]. For example, China is home to a quarter of the world's population; and as of 2016, China had 731 million Internet users (penetration rate: 53%) and 1.3 billion cell phone users (penetration rate: 95%; [25,26]). The Chinese government has been active in promoting the application of the Internet and mobile phones in health services delivery [27-29]. Studies on the use of mobile tools in special populations could be traced back to more than a decade ago, right from the 2000s [30-32]; recent literature indicates that both the Internet and mobile phones have been used in public health emergency responses [33], infectious disease surveillance [34], teleconsultation [35], and intervention delivery [36,37]. Despite the growing body of literature on the application of mobile tools, to date, no study exists on the digital divide in the Chinese population, especially in the older population.

With the aim to fill gaps in the literature, in this study we used data from the 2011 and 2013 waves of the China Health and Retirement Longitudinal Study (CHARLS) to examine the relationship between the digital divide and health disparities. We aim to answer the following research questions: (1) among the general middle-aged and elderly Chinese (45 years and older), what is the prevalence of access to mobile tools (Internet use and cell phone ownership)?; (2) what are the predictors of Internet use and mobile phone ownership at individual and neighborhood levels?; and (3) what is the relationship between health outcomes and access to mobile tools after controlling for potential confounders?

Methods

Data Source

The CHARLS is a nationally representative longitudinal survey of the middle-aged and elderly population (45+ years) in China. As detailed in previous reports [38-40], led by the Peking University in China via collaboration with the Oxford University in the United Kingdom and the University of Southern California in the United States, CHARLS was designed as a part of a set of international longitudinal aging surveys that include the Health and Retirement Study (HRS) in the United States, the Survey of Health and Retirement in Europe (SHARE), and similar longitudinal aging surveys in other countries. Following the protocols of the HRS, the CHARLS main questionnaire comprises seven modules, covering demographics, family background, health status, SES, and environment (community questionnaire and county-level policy questionnaire). All data were collected through face-to-face, computer-aided personal interviews (CAPI; [38-40]).

Sample Size

The national baseline was conducted from July 2011 to March 2012 and represented people aged 45 years and older, living in 150 counties in 28 provinces across China. A total of 150 county-level units were randomly selected using probability proportional to size (PPS) and stratified by region, urban/rural and county-level gross domestic product (GDP). Within each county-level unit, three village-level units (villages in rural areas and urban communities in urban areas) were randomly selected using PPS as primary sampling units (PSUs). Within each PSU, 80 dwellings were randomly selected from a complete list of dwelling units generated from a mapping or listing operation, using augmented Google Earth maps (Google Inc) along with considerable ground checking. In scenarios with more than one age-eligible household in a dwelling unit, one was randomly selected. From this sample for each PSU, the proportion of households with age-eligible members was determined, as was the proportion of empty residences. From these proportions and an assumed response rate, we selected households from our original PSU frame to obtain a target number of 24 age-eligible households per PSU. Thus, the final household sample size in a PSU depended on the PSU age-eligibility and empty residence rates. In each household, one person aged 45 years or older was randomly chosen as the main respondent, and the individual's spouse was automatically included. On the basis of this sampling procedure, 1 or 2 individuals in each household were interviewed depending on the marital status of the main respondent. The total sample size was 10,257 households and 10,481 individuals in the 2011 baseline. The sample size for the first follow-up in 2013 was 18,613 individuals [38-40].

Measures

Access to mobile technology was measured by Internet use and cell phone ownership. Internet use was measured by the question: "Have you accessed the Internet in the past month?" Cell phone use was measured by the question: "Do you own a cell phone?" Both measures were dichotomized into yes–no answers.

Demographic Characteristics

Demographic characteristics included age, gender, education, marital status, living arrangement, rural/urban residence (household registration or Hukou), employment status, and income. In this study, education was categorized into four groups: primary school or less, middle school, high school or vocational school, and some college or more. Living arrangement had three mutually exclusive categories: empty nester (living alone or with spouse), living with children, and living with someone other than children [41]. In terms of income, given that most rural residents have no regular income, expenditure is a better welfare measure than income in developing countries [42]. Thus in this study, we followed the other published studies from CHARLS and used per capita expenditure (PCE) of the past year in the household for income measure; PCE was log-transformed because of its skewness of distribution for analyzing its relationship with other variables [43-45].

Community Resources

Community resources were measured using the community-level survey conducted in 2011, which asked informed officials or personnel in the community about the characteristics of the communities in which the CHARLS respondents resided. Three measures of community resources were taken in this study: neighborhood amenities, health care facilities, and community organizations. Neighborhood amenities is a composite measure derived by summing the following services in the neighborhood: drinking water used (tap water, well, and river/spring), types of cooking fuel (gas, coal, and hay), waste disposal (moved away by truck, buried in village, burned, dumped into nearby river, or no management), and main toilet system (in-house, out-house, or open air; and for each type with or without flushing water). A composite score was created with a range of 0 to 14, with a higher score indicating more urbanized neighborhood amenities [41]. Health care facilities were measured by summing the number of health care facilities in the respondent's neighborhood, including the general hospital, specialized hospital, Chinese medicine hospital, community health center, township hospital, health care post, village clinic, private clinic, and pharmacy. A composite score ranging from 0 to 8 was created, with higher values indicating more health care resources [41]. Community resources is a composite measure derived by summing the following facilities in the community: basketball playground, swimming pool, outside exercise facilities, other outdoor sports facilities, room for card games and chess games, room for Ping-Pong, calligraphy and painting club, dancing team or other exercise clubs, other entertainment facilities, organizations for helping the elderly and the handicapped, activity center for residents, and an association for elders. A score ranging from 0 to 12 was created, with a higher value indicating more community resources [44].

Geographic Areas

Geographic areas were measured by the location where the survey was taken to capture the vast geographic differences in economic development and health care resources in China [45]. Three major regions of East China, Central China, and West China were included [46].

Health Outcomes

Health outcomes were measured by two indicators: self-reported health and having a disability. Such an approach allowed us to gauge older adults' perceptions of their health, both generally and specifically, in relation to performing daily activities [45]. Self-reported health was a subjective measure of one's health and was reported on the following scale: very good, good, fair, poor, or very poor. The responses were dichotomized to having good health (good or very good) and poor health (fair, not good, or poor) [43]. Having a disability was measured by two scales: the 6-item scale of activities of daily living (ADL) such as dressing and bathing and the 5-item scale of instrumental activities of daily living (IADL) such as preparing meals and taking medications. These 11 items were dichotomously coded (yes–no); no disability was defined as having no difficulty in all ADL or IADL items [43,46].

Data Analysis

First, we used the chi-square (for categorical variables) and *t*-test (for continuous variables) to examine the relationship between access to mobile tools and individual-level SES, community resources, health outcomes, and geographic regions. Second, because the CHARLS dataset has a natural hierarchical structure with individuals nested within the community and the aim of this paper was to analyze the effects of individual and community characteristics on the digital divide and health disparities, we used two-level mixed logistic regression, controlling for unobserved heterogeneity at the community and individual levels to explore the relationship between access to mobile tools and SES, community resources, and geographic regions [41,44,45]. Odds ratios (ORs) with CI were employed to depict the relationship between outcome variable and independent variable while controlling for other covariates. Finally, to examine the cross-sectional relationship between health outcomes and access to mobile tools, we made four separate multivariate logistic regression models for two dependent variables of health outcomes (self-reported health and having a disability) and two independent variables of interest (Internet access and mobile phone ownership) while controlling for potential confounders of SES, community resources, and geographic region. All analyses, both descriptive tables and regressions, were weighted using individual sampling weights with household and individual nonresponse adjusted. Sixteen variables with missing data were multiple imputed by the Windows software NORM [47]. All analyses were performed using Stata 13 (StataCorp).

Results

Participant Characteristics: Socioeconomic Status, Health Outcomes, and Community Resources

Of the 18,215 participants included in this study, 44% of the participants were from East China, 26% from Central China, and 30% from West China. The mean age was 61 years, and 51% of the participants were female. Approximately 62% of the participants had a primary school education or less, 22% had completed middle school, 13% had high school or vocational school education, and 3% had some college education or had completed college. More than 84% of the participants were married; about 59% of the participants lived alone or with a spouse, 6% lived with children, and 35% lived with others. About 70% of the participants were rural residents, and 39% were unemployed. The average PCE was 15,914.78 Chinese Yuan (or US $2316.80) per year. More than 26% of the participants reported having poor health and about 40% reported having a disability. The score for the average number of neighborhood amenities was 7.10 (range: 0-12); the mean number of health care facilities nearby was 1.37 (range: 0-6), and community resources was 4.40 (range: 0-14).

Binary Relationship of Access to Mobile Tools and SES, Health Outcomes, Community Resources, and Geographic Region

About 6.5% of the middle aged and elderly Chinese used the Internet, and 83% owned a mobile phone. As shown in Table 1, Internet access was associated with most of the variables of SES (except employment status), health outcomes, community resources, and geographic region. Specifically, significant age differences (54 years vs 61 years) existed between those who accessed Internet and those who did not. More males than females used the Internet (7.8% vs 5.3%), and Internet use was also associated with higher levels of education, being married, living with children, and urban residence; however, it was not related to employment status. Internet access was also associated with having good health and no disability. Three indicators of community resources were also significantly associated with Internet access, and so was geographic region. Similar patterns were observed in the binary relationship between mobile phone ownership and SES, health outcomes, community resources, and geographic region, with the exception of community organization.

Table 1. Relationship between the use of mobile technology (Internet use and mobile phone) and sociodemographic characteristics, community amenities, and geographic location (weighted).

Variables	Mean or category	Total	Internet use				Mobile phone use			
			Yes	No	χ^2(df) or *t* (df)	*P*	Yes	No	χ^2(df) or *t* (df)	*P*
Socioeconomic status										
Age, in years	Mean	60.54	54.07	60.99	−13.7 (17897)	<.001	58.88	68.70	−26.3 (17897)	<.001
Gender, n (%)	Female	9366 (51.42)	501 (5.35)	8865 (94.65)	43.2 (1)	.009	7696 (82.17)	1670 (17.83)	12.822 (1)	.04
	Male	8849 (48.58)	688 (7.78)	8160 (92.22)			7449 (84.18)	1400 (15.82)		
Education, n (%)	≤Primary school	11,224 (61.62)	156 (1.39)	11,068 (98.61)	2428.3 (3)	<.001	8910 (79.38)	2314 (20.62)	296.423 (3)	<.001
	≤Middle school	4029 (22.12)	268 (6.64)	3762 (93.36)			3589 (89.07)	440 (10.93)		
	≤High school/ Vocational school	2375 (13.04)	538 (22.67)	1837 (77.33)			2135 (89.89)	240 (10.11)		
	≥College	587 (3.22)	224 (38.25)	362 (61.75)			516 (87.92)	71 (12.08)		
Marital status, n (%)	Unmarried	2805 (15.40)	109 (3.88)	2696 (96.12)	37.1 (1)	.04	1779 (63.41)	1026 (36.59)	908.9 (1)	<.001
	Married	15,410 (84.60)	1077 (6.99)	14,333 (93.01)			13370 (86.76)	2040 (13.24)		
Living arrangement, n (%)	Empty nest (alone or with spouse)	10,730 (58.91)	931 (8.68)	9799 (91.32)	279.0 (2)	<.001	8690 (80.98)	2041 (19.02)	89.9 (2)	<.001
	Living with children	1186 (6.51)	112 (9.41)	1074 (90.59)			1041 (87.81)	145 (12.19)		
	Living with others	6299 (34.58)	144 (2.28)	6155 (97.72)			5418 (86.02)	881 (13.98)		
Rural-urban residence, n (%)	Urban	5554 (30.49)	984 (17.71)	4570 (82.29)	1615.5 (1)	<.001	4753 (85.59)	800 (14.41)	32.8 (1)	.008
	Rural	12,661 (69.51)	203 (1.60)	12,459 (98.40)			10396 (82.11)	2265 (17.89)		
Employment status, n (%)	Not	7095 (38.95)	480 (6.77)	6614 (93.23)	1.2 (1)	.68	5325 (75.05)	1770 (24.95)	537.2 (1)	<.001
	Yes	11,120 (61.05)	706 (6.35)	10,414 (93.65)			9824 (88.34)	1297 (11.66)		
Income (PCE[a]), in Chinese Yuan	Mean	15,914.78	33,953.60	14,657.84	7.9 (17897)	<.001	16,511.92	12,964.29	5.6 (17897)	<.001
Health outcomes										
Self-reported health, n (%)	Poor	4663 (25.60)	125 (2.68)	4538 (97.32)	148.8 (1)	<.001	3694 (79.22)	969 (20.78)	68.5	<.001
	Good	13,552 (74.40)	1061 (7.83)	12,491 (92.17)			11,455 (84.53)	2096 (15.47)		
Having a disability, n (%)	No	10,993 (60.35)	846 (7.70)	10,146 (92.3)	54.5 (1)	.003	9453 (85.99)	1540 (14.01)	110.6 (1)	<.001
	Yes	7222 (39.65)	352 (4.88)	6870 (95.12)			5779 (80.01)	1444 (19.99)		

Variables	Mean or category	Total	Internet use				Mobile phone use			
			Yes	No	χ^2(df) or *t* (df)	*P*	Yes	No	χ^2(df) or *t* (df)	*P*
Community resources										
Neighborhood amenities[b,c]	Mean	7.10	10.83	8.84	34.4 (17897)	<.001	7.15	6.9	1.8 (17897)	.08
Health care facilities[c,d]	Mean	1.37	1.79	1.34	5.3 (17897)	<.001	1.38	1.30	2.3 (17897)	.02
Community organizations[c,e]	Mean	4.40	7.14	4.21	14.2 (17897)	<.001	4.38	4.50	−0.7 (17897)	.47
Region										
East, n (%)		8058 (44.24)	728 (9.03)	7331 (90.97)	168.9 (2)	<.001	6713 (83.3)	1346 (16.70)	32.9 (1)	.003
Central, n (%)		4718 (25.90)	270 (5.73)	4447 (94.27)			3810 (80.76)	908 (19.24)		
West, n (%)		5439 (29.86)	188 (3.46)	5251 (96.54)			4626 (85.06)	813 (14.94)		

[a] PCE: Per capita expenditure.

[b] Neighborhood amenities comprised four variables (tap water, toilet, cooking fuel, and waste management); a range of 0 to 14, with higher value meaning higher coverage of modern amenities.

[c] Community/neighborhood data were not available for 2013; it was only available in 2011.

[d] Health facilities include six variables such as clinic, pharmacy, and hospital; with a range of 0 to 8.

[e] Community organizations include 14 variables such as having a senior activity room, having a community council, having a playground; a range of 0 to 12, with higher value meaning more community resources.

Multivariate Relationship of Access to Mobile Tools and SES, Community Resources, and Geographic Region

After controlling for potential confounders, Internet use was independently and significantly associated with the following predictors: age, gender, education level, marital status, living arrangement, rural/urban residency, income, neighborhood amenities, and geographic region, but it was not associated with employment status, health care facilities, and community organizations (see Table 2). Likewise, mobile phone ownership was independently and significantly associated with age, education level, marital status, living arrangement, rural/urban residency, income, and geographic region, but it was not associated with gender, employment status, neighborhood amenities, health care facilities, and community resources.

Multivariate Relationship of Access to Mobile Tools and Health Outcomes

Table 3 depicts the results of four models on the relationship of access to mobile tools and health outcomes. Internet access was significantly associated with self-reported health (adjusted odds ratio, aOR=1.73) but not related to disability while controlling for potential confounders. Mobile phone ownership was significantly associated with disability (aOR=0.843) but not with self-reported health. Other predictors of good self-reported health included gender, education, rural/urban residency, employment status, neighborhood amenities, and geographic region. Other predictors of having a disability included age, education, marital status, living arrangement, rural/urban residency, employment status, neighborhood amenities, community organization, and geographic region. We also analyzed the data by using health outcomes as continuous variables, and the results were similar.

Table 2. Estimates of fixed and random parameters from the multilevel mixed models of mobile technology use (Internet, mobile phone) on socioeconomic status, neighborhood amenities, and community resources (weighted).

Variables	Internet use		Mobile phone	
	aOR (95% CI)	*P*	aOR (95% CI)	*P*
Intercept	0.007 (0.001-0.044)	<.001	12.980 (4.733-35.600)	<.001
Individual-level variables				
Mean age, in years	0.912 (0.897-0.928)	<.001	0.928 (0.921-0.936)	<.001
Gender (Ref=female)	1.436 (1.207-1.708)	<.001	1.059 (0.989-1.134)	.10
Education (Ref=≤primary school)				
≤Middle school	3.951 (2.932-5.326)	<.001	1.238 (1.076-1.423)	.003
≤High school/Vocational school	9.409 (7.091-12.49)	<.001	1.336 (1.099-1.624)	.004
≥College	20.24 (13.90-29.45)	<.001	1.020 (0.673-1.547)	.93
Marital status (Ref=unmarried)	0.858 (0.592-1.243)	.42	1.948 (1.668-2.276)	<.001
Living arrangement (Ref=empty-nested)				
Living with children	0.888 (0.538-1.464)	.64	2.193 (1.634-2.943)	<.001
Living with others	0.886 (0.679-1.157)	.37	2.269 (1.970-2.614)	<.001
Rural-urban residence (Ref=urban)	0.368 (0.279-0.485)	<.001	0.553 (0.455-0.672)	<.001
Employment status (Ref=unemployed)	1.047 (0.823-1.332)	.71	1.505 (1.328-1.706)	<.001
Log income, mean (per capita expenditure)	1.593 (1.407-1.802)	<.001	1.441 (1.340-1.549)	<.001
Community-level variables				
Community resources				
Neighborhood amenities, mean	1.188 (1.127-1.253)	<.001	1.008 (0.977-1.040)	.62
Health facilities, mean	1.037 (0.924-1.164)	.54	0.995 (0.914-1.082)	.89
Community organization, mean	1.018 (0.976-1.063)	.41	0.960 (0.926-0.994)	.02
Region (Ref=east)				
Central	1.328 (0.965-1.828)	.08	0.875 (0.707-1.083)	.22
West	0.755 (0.534-1.069)	.11	1.259 (1.020-1.553)	.03
Random effect variance	1.734 (1.420-2.116)	<.001	1.741 (1.527-1.984)	<.001

Table 3. Estimates of fixed and random parameters from the multilevel mixed models of health outcomes and mobile technology use (weighted).

Variables	Internet access as primary predictor				Mobile phone ownership as primary predictor			
	Self-reported health		Having a disability		Self-reported health		Having a disability	
	aOR (95% CI)	*P*	aOR (95% CI)	*P*	aOR (95% CI)	*P*	aOR (95% CI)	*P*
Intercept	2.421 (1.330-4.406)	.004	0.193 (0.114-0.328)	<.001	2.298 (1.245-4.240)	.008	0.229 (0.133-0.392)	<.001
Internet use (Ref=no Internet use)	1.727 (1.327-2.246)	<.001	1.138 (0.945-1.371)	.17	__[a]		__[a]	
Own mobile phone (Ref=no mobile phone)	__[b]		__[b]		1.076 (0.968-1.196)	.17	0.843 (0.763-0.931)	<.001
Individual-level variables								
Age	0.988 (0.983-0.992)	<.001	1.022 (1.018-1.026)	<.001	0.988 (0.983-0.992)	<.001	1.020 (1.015-1.024)	<.001
Gender (Ref=female)	1.373 (1.278-1.475)	<.001	1.010 (0.944-1.080)	.77	1.374 (1.279-1.476)	<.001	1.012 (0.946-1.082)	.74
Education (Ref=≤primary school)								
≤Middle school	1.121 (1.023-1.228)	.01	0.874 (0.808-0.945)	<.001	1.129 (1.031-1.236)	.009	0.878 (0.812-0.949)	.001
≤High school/Vocational school	1.282 (1.110-1.481)	<.001	0.890 (0.792-1.001)	.05	1.333 (1.155-1.538)	<.001	0.905 (0.807-1.016)	.09
≥College	1.486 (1.099-2.010)	.01	0.887 (0.675-1.166)	.39	1.659 (1.228-2.239)	<.001	0.917 (0.703-1.196)	.52
Marital status (Ref=unmarried)	0.913 (0.816-1.022)	.11	0.827 (0.742-0.922)	<.001	0.902 (0.805-1.010)	.07	0.842 (0.755-0.938)	.002
Living arrangement (Ref=empty nested)								
Living with children	1.047 (0.875-1.252)	.62	1.114 (0.967-1.285)	.14	1.038 (0.866-1.243)	.07	1.129 (0.978-1.304)	.09
Living with others	1.139 (1.054-1.232)	.001	0.951 (0.885-1.023)	.02	1.128 (1.041-1.221)	.003	0.965 (0.897-1.038)	.34
Rural/urban residence (Ref=urban)	0.731 (0.648-0.824)	<.001	1.295 (1.167-1.437)	<.001	0.722 (0.640-0.814)	<.001	1.273 (1.148-1.413)	<.001
Employment status (Ref=unemployed)	2.239 (2.044-2.454)	<.001	0.619 (0.568-0.674)	<.001	2.235 (2.040-2.449)	<.001	0.623 (0.572-0.679)	<.001
Log income, mean (per capita expenditure)	1.01 (0.969-1.051)	.64	1.028 (0.991-1.067)	.13	1.011 (0.971-1.052)	.60	1.037 (0.999-1.076)	.05
Community-level variables								
Community resources								
Neighborhood amenities, mean	1.074 (1.055-1.093)	<.001	0.976 (0.961-0.990)	.001	1.076 (1.057-1.095)	<.001	0.976 (0.962-0.991)	.002
Health facilities, mean	0.997 (0.946-1.051)	.90	1.008 (0.964-1.055)	.72	0.997 (0.946-1.051)	.91	1.008 (0.964-1.055)	.72
Community organization, mean	1.017 (0.998-1.035)	.07	0.982 (0.967-0.998)	.03	1.017 (0.999-1.036)	.06	0.982 (0.966-0.998)	.02
Region (Ref=east)								
Central	0.878 (0.772-0.999)	.05	1.272 (1.137-1.422)	<.001	0.883 (0.775-1.005)	.06	1.270 (1.135-1.422)	<.001
West	0.775 (0.687-0.875)	<.001	1.541 (1.382-1.718)	<.001	0.772 (0.684-0.871)	<.001	1.547 (1.387-1.726)	<.001
Random effect variance	1.120 (1.078-1.163)	<.001	1.102 (1.070-1.135)	<.001	1.120 (1.078-1.163)	<.001	1.104 (1.072-1.138)	<.001

[a] Mobile phone ownership as primary predictor, the parameters of Internet use are missing.

[b] Internet access as primary predictor, the parameters of mobile phone ownership are missing.

Discussion

Principal Findings

Our data analysis revealed that only a small percentage (6.5%) of middle-aged and elderly Chinese participants accessed the Internet, but a high proportion (83%) of the participants owned a mobile phone. The rate of access to the Internet was much lower than that of the official report of Internet access in the general population (53%; [25]). Such a discrepancy might be because of three possible reasons. First, our measure of Internet access was based on the question—"Have you accessed Internet in the past month?," but the other surveys on Internet use typically measured lifetime use. Second, many older adults interpret "accessed Internet" to mean going online via a computer only; many have used mobile phones for Web-based activities but did not report so. Third, prior studies have relied on voluntary convenience sampling, and middle-aged and elderly people and rural residents were less likely to be included in the survey; however, the rate of mobile phone ownership was comparable with that of general population (95%; [26]).

Our findings corroborated existing literature on the relationship of SES and digital divide by adding new evidence from middle-aged and elderly Chinese. Similar to other countries, access to mobile tools was associated with younger age, higher level of education, higher income, and urban residence [4,8,11,12,17,20,22]. Our data also suggested different sets of SES predictors for Internet access and mobile phone ownership; for example, significantly fewer women than men accessed Internet, but gender was not a significant predictor of mobile phone ownership. The high rate of mobile phone ownership in China may provide an equal opportunity for women to access information [48]. By contrast, people who were married or living with children or others were more likely to own mobile phones compared with people who are single or living alone, but Internet access was not associated with marital status or living arrangement. This may suggest that the mobile phone has instead become an important communication tool for people living with families or others.

Our study also examined the effect of community-level SES on access to mobile tools. We found that Internet access was strongly associated with neighborhood amenities (drinking water, toilet, etc) but not health care facilities or community organization; mobile phone ownership was not associated with any of the three measures of community resources. Neighborhood amenities were a good indicator of urbanization [43], and their association with Internet access suggests that Internet access may be considered a kind of community resource. Because urban areas are better equipped with broadband access, urban residents are more likely to access the Internet. The lack of a relationship between mobile phone ownership and community resources suggests that mobile phone as a portable and personal communication tool has wider accessibility, as it is less likely to be restricted by community-level facilities or resources.

In the analysis on the relationship of the digital divide and health disparities, our data showed that self-reported health was significantly associated with Internet use, and mobile phone ownership was significantly associated with not having a disability. Such findings were consistent with the literature on the relationship of Internet access and status [3,13]. As documented in the literature and also described above, Internet access and mobile phone ownership were important indicators of SES at individual and community levels; therefore, the relationship between Internet access and health status was de facto the reciprocal relationship of SES and health outcomes [4,10,17].

We have also observed significant health disparities in our participants; health disparities were predicted by SES, rural/urban residence, community resources, and geographic region, which were consistent with prior CHARLS studies [41,43-46]. Such disparities reflect the inequality of resource allocation and economic development in rural and urban areas and across regions.

The following policy implications are associated with the forgoing empirical findings on the digital divide and health disparities in China. First, very few middle-aged and elderly Chinese use the Internet, which is a strong predictor of SES and health outcomes. Because community SES has been recognized as a strong predictor of individual health [16], building community resources has been advocated as an important strategy to improve health [43,44]. China promotes the "Internet of Things" and "intelligent hospitals" [27-29], so improving access to mobile tools, especially the Internet for underserved communities and in underdeveloped areas, would potentially yield significant improvement in health outcomes.

Second, a majority of older Chinese have owned a mobile phone. The high ownership rate of the mobile phone suggests that it could become a tool that transcends social classes and reaches the vulnerable and underserved. If welfare covers food and rent as necessities of life, vouchers for mobile phone subscriptions could also be considered so that the elderly and people with a disability or those who are living alone could have access to this basic communication tool [8,10]. Existing literature has shown that closing the digital gap is conducive to bridging health [49].

Third, as more people are connected with mobile phones, such ubiquitous access can be maximized upon for empowerment and health services delivery. Researchers have documented initial evidence of the efficacy of mobile phone–delivered health intervention [50]; however, most of these mobile health (mHealth) interventions were conducted in Western countries with limited data from developing countries such as China, despite a high ownership rate of mobile phones. In recent years, some scientists have piloted mobile-based interventions [36,37,51-53]. For example, a recent study showed that text messages could effectively promote smoking cessation in Chinese adults [37]. Such endeavors would be especially beneficial for the elderly Chinese, given that China is aging rapidly, thanks to its improved life expectancy and the decline in fertility. The three-decade one-child policy has dramatically affected the elderly care model in China, and the current resources cannot keep up with the rapidly growing aging population [54]. Some experts have called for innovative mHealth solutions for chronic condition management and elderly care in China [55]. Widely accessible mobile phones and continuing penetration of Internet access may be a part of the solutions.

Strengths and Limitations

Our study has the following strengths. First, it was based on a national probability sample of older adults in China and had a large sample size; therefore, our findings can be generalized to other older adults in China. Second, the SES measures included variables at individual, household, and community levels, thus giving us a comprehensive measurement of SES. Third, our health outcomes were measured by both self-reported health and disability using two scales and 11 items relevant to older adults in China.

We also note the following limitations in our study. First, the study design was cross-sectional in nature, and we could infer no causal relationship. In addition, the community resources were collected in 2011 (data was not available in 2013), but health outcomes, access to mobile tools, and other covariates were collected in 2013. The relationship of community resources and access to the Internet and health outcomes may be predictive. Second, the CHARLS includes only 2 simple questions on access to mobile tools, so we could measure only Internet use and mobile phone ownership. Other important aspects of mobile tool use, including the length of use, frequency of use, purpose of use, and whether they use smartphones, were missing in this study; however, a recent survey has shown that smartphone users in China accounted for 53% of all mobile phone users in 2016 [56]. Whether and how people use mobile tools for health purposes affects the relationship of the digital divide and health disparities [57]; therefore, we call for more data on access to and usage of mobile tools. Third, our measures of health outcomes were also limited to two indicators. Furthermore, the dichotomous nature of the health outcome variables may limit our analysis of the relationship of the digital divide and health disparities.

Conclusions

To conclude, to the best of our knowledge, this study is the first on the digital divide and health disparities in China. Our study advanced the literature by providing data on the relationship of SES and the digital divide by embedding individual characteristics in community resources. A low rate of Internet access and its strong relationship with neighborhood amenities and health outcomes suggest that Internet access may be an important indicator of individual- and community-level SES, and more Internet access may result in upgrades in individual SES and community infrastructure. By contrast, the high ownership rate of mobile phone and its lack of relationship with community resources suggest that the mobile phone may transcend social classes and become an ordinary personal item. The high ownership of mobile phones presents an enormous potential to empower and serve older Chinese, who face mounting challenges in care in a rapidly aging society. We call for more studies on the use of mobile tools and its relationship with health disparities in China.

Conflicts of Interest

None declared.

References

1. Chou WS, Liu B, Post S, Hesse B. Health-related internet use among cancer survivors: data from the health information national trends survey, 2003-2008. J Cancer Surviv 2011 Sep;5(3):263-270. [doi: 10.1007/s11764-011-0179-5] [Medline: 21505861]
2. McCully SN, Don BP, Updegraff JA. Using the Internet to help with diet, weight, and physical activity: results from the health information national trends survey (HINTS). J Med Internet Res 2013 Aug 01;15(8):e148 [FREE Full text] [doi: 10.2196/jmir.2612] [Medline: 23906945]
3. DiMaggio P, Hargittai E, Celeste C, Shafer S. From unequal access to differentiated use: a literature review and agenda for research on digital inequality. In: Neckerman K, editor. Social Inequality. New York: Russell Sage Foundation; 2004:355-400.
4. Hale T. Is there such a thing as an online health lifestyle? Information, Communication & Society 2013 May;16(4):501-518. [doi: 10.1080/1369118X.2013.777759]
5. NTIA. Falling through the net: defining the digital divide URL: https://www.ntia.doc.gov/report/1999/falling-through-net-defining-digital-divide [WebCite Cache ID 6rti9Aw45]
6. 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]
7. Robinson L, Cotten S, Ono H, Quan-Haase A, Mesch G, Chen W, et al. Digital inequalities and why they matter. Information, Communication & Society 2015 Mar 16;18(5):569-582. [doi: 10.1080/1369118X.2015.1012532]
8. 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 02;15(5):e93 [FREE Full text] [doi: 10.2196/jmir.2645] [Medline: 23639979]

9. Kontos E, Emmons KM, Puleo E, Viswanath K. Communication inequalities and public health implications of adult social networking site use in the United States. J Health Commun 2010;15(Suppl 3):216-235 [FREE Full text] [doi: 10.1080/10810730.2010.522689] [Medline: 21154095]
10. Viswanath K, Kreuter MW. Health disparities, communication inequalities and e-health: a commentary. Am J Prev Med 2007;32(5 Suppl):S131-S133. [doi: 10.1016/j.amepre.2007.02.012]
11. Hale T, Goldner M, Stern M, Drentea P, Cotten S. Patterns of online health searching 2002–2010: implications for social capital, health disparities and the de-professionalization of medical knowledge. In: Technology, Communication, Disparities and Government Options in Health and Health Care Services. Bingley, West Yorkshire: Emerald Group Publishing Limited; 2014:35-60.
12. Hale T, Cotten S, Drentea P, Goldner M. Rural-urban differences in general and health-related internet use. Am Behav Sci 2010;53(9):1304-1325. [doi: 10.1177/0002764210361685]
13. Zillien N, Hargittai E. Digital distinction: status ‐ specific types of internet usage. Soc Sci Q 2009;90(2):274-291. [doi: 10.1111/j.1540-6237.2009.00617.x]
14. Link B, Phelan J. Social conditions as fundamental causes of disease. J Health Soc Behav 1995;Spec No:80-94. [Medline: 7560851]
15. Diez RA. Investigating neighborhood and area effects on health. Am J Public Health 2001 Nov;91(11):1783-1789. [Medline: 11684601]
16. Kawachi I, Berkman L. Neighborhoods and health. New York, NY: Oxford University Press; 2003.
17. Choi N. Relationship between health service use and health information technology use among older adults: analysis of the US National Health Interview Survey. J Med Internet Res 2011 Apr 20;13(2):e33 [FREE Full text] [doi: 10.2196/jmir.1753] [Medline: 21752784]
18. Gracia E, Herrero J. Internet use and self-rated health among older people: a national survey. J Med Internet Res 2009 Dec 02;11(4):e49 [FREE Full text] [doi: 10.2196/jmir.1311] [Medline: 19955041]
19. Cotten S, Ford G, Ford S, Hale T. Internet use and depression among retired older adults in the United States: a longitudinal analysis. J Gerontol B Psychol Sci Soc Sci 2014 Sep;69(5):763-771. [doi: 10.1093/geronb/gbu018] [Medline: 24671896]
20. Czaja SJ, Sharit J, Lee CC, Nair SN, Hernández MA, Arana N, et al. Factors influencing use of an e-health website in a community sample of older adults. J Am Med Inform Assoc 2013;20(2):277-284 [FREE Full text] [doi: 10.1136/amiajnl-2012-000876] [Medline: 22802269]
21. Leist A. Social media use of older adults: a mini-review. Gerontology 2013;59(4):378-384. [doi: 10.1159/000346818] [Medline: 23594915]
22. 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 2016 Aug 24:- Epub ahead of print. [doi: 10.1093/geronb/gbw100] [Medline: 27558403]
23. Fuchs C, Horak E. Africa and the digital divide. Telemat Inform 2008 May;25(2):99-116. [doi: 10.1016/j.tele.2006.06.004]
24. Venkatesh V, Sykes T. Digital divide initiative success in developing countries: a longitudinal field study in a village in India. Inf Syst Res 2013 Jun;24(2):239-260. [doi: 10.1287/isre.1110.0409]
25. China Internet Watch. Whitepaper: China internet statistics 2017 URL: https://www.chinainternetwatch.com/whitepaper/china-internet-statistics/ [accessed 2017-03-29] [WebCite Cache ID 6pKVCEEeq]
26. China Internet Watch. 95 out of 100 Chinese own a mobile phone in 2015 URL: https://www.chinainternetwatch.com/17105/china-mobile-phone-users-2015/ [accessed 2017-03-29] [WebCite Cache ID 6pKVNInPS]
27. Hvistendahl M. Information technology. China pushes the 'Internet of Things'. Science 2012 Jun 08;336(6086):1223. [doi: 10.1126/science.336.6086.1223] [Medline: 22679075]
28. Tu J, Wang C, Wu S. The internet hospital: an emerging innovation in China. Lancet Glob Health 2015 Aug;3(8):e445-e446 [FREE Full text] [doi: 10.1016/S2214-109X(15)00042-X] [Medline: 26187488]
29. Zheng X, Rodríguez-Monroy C. The development of intelligent healthcare in China. Telemed J E Health 2015 May;21(5):443-448. [doi: 10.1089/tmj.2014.0102] [Medline: 25671730]
30. Hong Y, Li X, Mao R, Stanton B. Internet use among Chinese college students: implications for sex education and HIV prevention. Cyberpsychol Behav 2007 Apr;10(2):161-169. [doi: 10.1089/cpb.2006.9973] [Medline: 17474831]
31. Hong Y, Li X, Fang X, Lin X, Zhang C. Internet use among female sex workers in China: implications for HIV/STI prevention. AIDS Behav 2011 Feb;15(2):273-282. [doi: 10.1007/s10461-010-9846-1] [Medline: 21082341]
32. Corpman D. Mobile health in China: a review of research and programs in medical care, health education, and public health. J Health Commun 2013;18(11):1345-1367. [doi: 10.1080/10810730.2013.778370] [Medline: 23915211]
33. Yang C, Yang J, Luo X, Gong P. Use of mobile phones in an emergency reporting system for infectious disease surveillance after the Sichuan earthquake in China. Bull World Health Organ 2009 Aug;87(8):619-623 [FREE Full text] [Medline: 19705013]
34. Wang L, Liu X, Huang F, Hennig C, Uplekar M, Jiang S. Engaging hospitals to meet tuberculosis control targets in China: using the internet as a tool to put policy into practice. Bull World Health Organ 2010 Dec 01;88(12):937-942 [FREE Full text] [doi: 10.2471/BLT.09.071753] [Medline: 21124719]

35. Wang TT, Li JM, Zhu CR, Hong Z, An DM, Yang HY, et al. Assessment of utilization and cost-effectiveness of telemedicine program in western regions of China: a 12-year study of 249 hospitals across 112 cities. Telemed J E Health 2016 Nov;22(11):909-920. [doi: 10.1089/tmj.2015.0213] [Medline: 27314300]
36. Chen S, Gong E, Kazi DS, Gates AB, Karaye KM, Girerd N, et al. Development of a mobile phone-based intervention to improve adherence to secondary prevention of coronary heart disease in China. J Med Eng Technol 2016;40(7-8):372-382. [doi: 10.1080/03091902.2016.1213904] [Medline: 27686107]
37. Liao Y, Wu Q, Tang J, Zhang F, Wang X, Qi C, et al. The efficacy of mobile phone-based text message interventions ('Happy Quit') for smoking cessation in China. BMC Public Health 2016 Aug 19;16(1):833 [FREE Full text] [doi: 10.1186/s12889-016-3528-5] [Medline: 27543164]
38. Zhao Y, Strauss J, Yang G, Giles J, Hu P, Hu Y. CCER. 2013. China health and retirement longitudinal study-2011-2012 URL: http://charls.ccer.edu.cn/uploads/document/2011-charls-wave1/application/CHARLS_users__guide_national_baseline_survey-English_20130407.pdf [accessed 2017-03-29] [WebCite Cache ID 6pKVS0jqS]
39. Zhao Y, Hu Y, Smith JP, Strauss J, Yang G. Cohort profile: the China health and retirement longitudinal study (CHARLS). Int J Epidemiol 2014 Feb;43(1):61-68 [FREE Full text] [doi: 10.1093/ije/dys203] [Medline: 23243115]
40. Zhao Y, Crimmins E, Hu P, Hu Y, Ge T, Kim J, et al. CCER. China health and retirement longitudinal study-2011-2012 blood data users guide, 2014 URL: http://charls.ccer.edu.cn/uploads/document/2011-charls%20wave1/application/blood_user_guide_en_20140429.pdf [accessed 2017-03-29] [WebCite Cache ID 6pKVUhaf8]
41. Li LW, Liu J, Zhang Z, Xu H. Late-life depression in Rural China: do village infrastructure and availability of community resources matter? Int J Geriatr Psychiatry 2015 Jul;30(7):729-736 [FREE Full text] [doi: 10.1002/gps.4217] [Medline: 25333218]
42. Strass J, Thomas D. Health over the life course. In: Handbook of Development Economics, volume 4. Amsterdam: Elsevier BV; 2008:3375-3474.
43. Smith JP, Tian M, Zhao Y. Community effects on elderly health: evidence from CHARLS national baseline. J Econ Ageing 2013 Nov;1-2:50-59 [FREE Full text] [doi: 10.1016/j.jeoa.2013.07.001] [Medline: 24386630]
44. Shen Y. Community building and mental health in mid-life and older life: evidence from China. Soc Sci Med 2014 Apr;107:209-216. [doi: 10.1016/j.socscimed.2013.12.023] [Medline: 24485943]
45. Evandrou M, Falkingham J, Feng Z, Vlachantoni A. Individual and province inequalities in health among older people in China: evidence and policy implications. Health Place 2014 Nov;30:134-144 [FREE Full text] [doi: 10.1016/j.healthplace.2014.08.009] [Medline: 25262491]
46. Lei X, Strauss J, Tian M, Zhao Y. Living arrangements of the elderly in China: evidence from the CHARLS national baseline. China economic journal 2015;8(3):191-214. [doi: 10.1080/17538963.2015.1102473]
47. Schafer JL, Olsen MK. Multiple imputation for multivariate missing-data problems: a data analyst's perspective. Multivariate Behav Res 1998 Oct 01;33(4):545-571. [doi: 10.1207/s15327906mbr3304_5] [Medline: 26753828]
48. Chen D, Derek H. SSRN. 2004. Gender equality and economic development: the role for information and communication technologies URL: https://ssrn.com/abstract=610329 [accessed 2017-03-29] [WebCite Cache ID 6pKVXyvjR]
49. Viswanath K. Cyberinfrastructure: an extraordinary opportunity to bridge health and communication inequalities? Am J Prev Med 2011 May;40(5 Suppl 2):S245-S248. [doi: 10.1016/j.amepre.2011.02.005] [Medline: 21521602]
50. Hall A, Cole-Lewis H, Bernhardt J. Mobile text messaging for health: a systematic review of reviews. Annu Rev Public Health 2015 Mar 18;36:393-415 [FREE Full text] [doi: 10.1146/annurev-publhealth-031914-122855] [Medline: 25785892]
51. Li Y, Wang W, van VM, Chen L, Car J, Rudan I, et al. Text messaging data collection for monitoring an infant feeding intervention program in rural China: feasibility study. J Med Internet Res 2013 Dec 04;15(12):e269 [FREE Full text] [doi: 10.2196/jmir.2906] [Medline: 24305514]
52. Chen Y, Yang K, Jing T, Tian J, Shen X, Xie C, et al. Use of text messages to communicate clinical recommendations to health workers in rural China: a cluster-randomized trial. Bull World Health Organ 2014 Jul 01;92(7):474-481 [FREE Full text] [doi: 10.2471/BLT.13.127076] [Medline: 25110372]
53. van VM, Li Y, Wang W, Chen L, Du X, Wu Q, et al. Prevalence of mobile phones and factors influencing usage by caregivers of young children in daily life and for health care in rural China: a mixed methods study. PLoS One 2015;10(3):e0116216.
54. Zhang Y, Goza F. Who will care for the elderly in China?: a review of the problems caused by China's one-child policy and their potential solutions. J Aging Stud 2006;20(2):151-164. [doi: 10.1016/j.jaging.2005.07.002]
55. Sun J, Guo Y, Wang X, Zeng Q. mHealth for aging China: opportunities and challenges. Aging Dis 2016 Jan;7(1):53-67 [FREE Full text] [doi: 10.14336/AD.2015.1011] [Medline: 26816664]
56. Statista. Number of smartphone users in China from 2013 to 2021 URL: https://www.statista.com/statistics/467160/forecast-of-smartphone-users-in-china/ [accessed 2017-07-12] [WebCite Cache ID 6rtfQDvUW]
57. Viswanath K, Finnegan J, Gollust S. Communication and health behavior in a changing media environment. In: Health Behavior: Theory, Research, and Practice, 5th Edition. San Francisco, CA: Jossey-Bass; 2015:327-348.

Abbreviations

ADL:	6-item scale of activities of daily living
aOR:	adjusted odds ratio
CAPI:	computer-aided personal interviews
CHARLS:	China Health and Retirement Longitudinal Study
eHealth:	electronic health
GDP:	gross domestic product
HRS:	Health and Retirement Study
IADL:	5-item scale of instrumental activities of daily living
mHealth:	mobile health
OR:	odds ratio
PCE:	per capita expenditure
PPS:	probability proportional to size
SES:	socioeconomic status
SHARE:	Survey of Health and Retirement in Europe

Edited by G Eysenbach; submitted 29.03.17; peer-reviewed by S McCully, M Almalki; comments to author 28.06.17; revised version received 12.07.17; accepted 12.07.17; published 11.09.17.

Please cite as:
Hong YA, Zhou Z, Fang Y, Shi L
The Digital Divide and Health Disparities in China: Evidence From a National Survey and Policy Implications
J Med Internet Res 2017;19(9):e317
URL: http://www.jmir.org/2017/9/e317/
doi:10.2196/jmir.7786
PMID:28893724

Original Paper

Internet Searches and Their Relationship to Cognitive Function in Older Adults: Cross-Sectional Analysis

Johanna Austin[1], PhD; Kristy Hollingshead[2], PhD; Jeffrey Kaye[1], MD

[1]Department of Neurology, Oregon Health & Science University, Portland, OR, United States

[2]Florida Institute for Human & Machine Cognition, Pensacola, FL, United States

Corresponding Author:
Johanna Austin, PhD
Department of Neurology
Oregon Health & Science University
3303 SW Bond Ave
Portland, OR, 97239
United States
Phone: 1 503 418 9328
Fax: 1 503 418 9311
Email: peterjo@ohsu.edu

Abstract

Background: Alzheimer disease (AD) is a very challenging experience for all those affected. Unfortunately, detection of Alzheimer disease in its early stages when clinical treatments may be most effective is challenging, as the clinical evaluations are time-consuming and costly. Recent studies have demonstrated a close relationship between cognitive function and everyday behavior, an avenue of research that holds great promise for the early detection of cognitive decline. One area of behavior that changes with cognitive decline is language use. Multiple groups have demonstrated a close relationship between cognitive function and vocabulary size, verbal fluency, and semantic ability, using conventional in-person cognitive testing. An alternative to this approach which is inherently ecologically valid may be to take advantage of automated computer monitoring software to continually capture and analyze language use while on the computer.

Objective: The aim of this study was to understand the relationship between Internet searches as a measure of language and cognitive function in older adults. We hypothesize that individuals with poorer cognitive function will search using fewer unique terms, employ shorter words, and use less obscure words in their searches.

Methods: Computer monitoring software (WorkTime, Nestersoft Inc) was used to continuously track the terms people entered while conducting searches in Google, Yahoo, Bing, and Ask.com. For all searches, punctuation, accents, and non-ASCII characters were removed, and the resulting search terms were spell-checked before any analysis. Cognitive function was evaluated as a z-normalized summary score capturing five unique cognitive domains. Linear regression was used to determine the relationship between cognitive function and Internet searches by controlling for variables such as age, sex, and education.

Results: Over a 6-month monitoring period, 42 participants (mean age 81 years [SD 10.5], 83% [35/42] female) conducted 2915 searches using these top search engines. Participants averaged 3.08 words per search (SD 1.6) and 5.77 letters per word (SD 2.2). Individuals with higher cognitive function used more unique terms per search (beta=.39, *P*=.002) and employed less common terms in their searches (beta=1.39, *P*=.02). Cognitive function was not significantly associated with the length of the words used in the searches.

Conclusions: These results suggest that early decline in cognitive function may be detected from the terms people search for when they use the Internet. By continuously tracking basic aspects of Internet search terms, it may be possible to detect cognitive decline earlier than currently possible, thereby enabling proactive treatment and intervention.

(J Med Internet Res 2017;19(9):e307) doi:10.2196/jmir.7671

KEYWORDS

Internet; geriatrics; cognition; executive function

Introduction

One in 9 adults over the age of 65 has a diagnosis of Alzheimer disease (AD), the 6th leading cause of death in the United States [1]. Due to the significant public health importance of this disease, many clinical trials have been performed in search of an effective treatment. However, from 2002-2012, the clinical trial success of AD drugs advancing to market was 1 out of 244 tested compounds [2], leading some to call this AD's "lost decade" [3]. This may in part be due to challenges in early diagnosis of AD. AD is marked by an insidious onset and gradual, subtle decline of cognitive function. By the time cognitive and functional symptoms are detected through clinical assessment, disease progression may already be too advanced for treatment to be most effective. Thus, a major focus of AD research is currently directed to early detection and prevention of the disease [4].

AD is likely the result of the progressive accrual of neuropathological lesions in the brain which ultimately affect multiple aspects of cognition and behavior [5]. Abilities that may be affected by this process are evident in a number of domains that are typically assessed in a clinical evaluation, such as language processing, motor function, and executive functioning. However, assessment during clinic visits may not allow for the detection of subtle changes in real-world behavior that are associated with changes in cognitive function. Furthermore, episodic assessment provides only a snapshot of cognitive function and makes detection of crucial changes difficult. Given that many data points are needed to detect a statistically significant change, this episodic assessment paradigm requires multiple years of data. In contrast, recent work by Dodge et al demonstrated that by collecting more frequent (eg, daily) measures of key variables of interest, it is possible to detect trends and changes in variables over a much shorter period (eg, months instead of years), enabling more sensitive detection of the earliest stages of decline [6]. As subtle changes in everyday function take years to evolve and manifest dementia, the ability to detect these changes through continuous in-home monitoring of everyday behavior and activity holds great promise for early detection of AD [7-10].

One key behavior that may enable real-world identification of daily function is computer use. Engaging with a computer requires motor function to operate the keyboard and move the mouse; language processing to comprehend, select, retrieve, and generate appropriate words; and executive function to plan, inhibit, focus, and shift attention in meaningful and efficient ways. Thus, the way an individual interacts with their computer represents a rich and relatively untapped means of assessing everyday cognition especially among older adults who are at risk of cognitive decline. Notably, although adoption of computers among older adults has lagged younger generations [11], as the baby boomer generation ages, the number of computer-savvy older adults is likely to increase dramatically [12]. Thus, an understanding of the relationship between regular computer use and cognitive function will become highly valuable, especially as 80% of older adults who use the Internet go online at least 3 times per week [12]. This computer monitoring approach has begun to be successfully carried out in older adults by assessing a number of computer use metrics for indexing cognitive change. These include general time-use metrics [13] as well as more specific operational aspects, for example, how a person completes an online task or operates a keyboard [14,15] or mouse [7].

One area of particular potential in this regard is inferring aspects of cognitive function and, more specifically, language function through the terms people search for on the Internet. Previous studies on language analysis (not typically done using computer-based monitoring) have indicated that certain key aspects of language decline in neurocognitive disorders such as AD [16]. These language changes are likely due to disruption of regions of the brain responsible for production and encoding of language [17-20]. Importantly, changes in language function have been observed before the clinical diagnosis of manifest dementia [21]. Aspects of language that decline include semantic fluency, picture naming, and phonetic fluency [22]. More recently, researchers have begun using automatic speech detection or recognition (ASR) systems to analyze not only word use but also pauses and speech tempo [23] and have demonstrated that these key aspects of language are sensitive to cognitive decline before other clinical tests may detect the disease [24,25].

All of this recent research has relied on elicited speech or language using a formal testing paradigm. This study bridges the gap between the previous work on inferring cognitive function from computer use and the work regarding the effects of cognitive decline on language by using continuous computer monitoring software to collect samples of language from the terms people search for on the Internet. Search term language differs from spoken language in that it is frequently goal oriented, may use only key words rather than complete sentences, and may also use terms not used in a spoken language such as "df" instead of "definition." Although a large body of literature has focused on how older adults search the Internet to find health information [26-28], relatively few studies have investigated general aspects of search term language or how they may relate to cognitive function.

The focus of this paper was to determine whether early language changes can be detected from the way people search the Internet during routine, everyday use. In particular, we hypothesize that individuals with more impaired cognitive function will (1) employ fewer unique search terms per search, (2) employ shorter words in their searches, and (3) use less obscure search terms, where obscurity is defined as the inverse of the frequency of searching for a given term across all subjects.

Methods

Participants

The participants for this study were recruited from two ongoing projects: the Life Laboratory cohort (Oregon Health & Science University (OHSU) IRB #2765) and the Ambient Independence Measures (AIMs) cohort (OHSU IRB #9944). The focus of these studies is to understand the relationship between daily behavior and health in older adults using home embedded sensing and computing ("smart home") technologies [9].

Eligibility criteria for both studies included living alone and independently in a house or apartment larger than a studio, a minimum score of 25 on the Mini-Mental State Examination, and a maximum score of 0.5 on the Clinical Dementia Rating scale (not demented). Participants were also required to live independently without the need for in-home nursing care or help with daily activities. The minimum age for participation in the Life Laboratory study is 62 years, whereas that for the AIMS cohort is 70 years. Enrollment for the Life Laboratory study began in 2007 and continues on a rolling basis. Enrollment for the AIMs study began in 2014 and closed in April 2016.

In 2015, all active participants from both projects who indicated they used a computer were asked to participate in this additional computer use monitoring study. Participants who agreed to have their computer use monitored received software (WorkTime, Nestersoft Inc) on their personal computers, which records all activities performed on the computer, including terms searched for in search engines such as Google, Yahoo, Bing, and Ask.com. All subjects signed informed consent before participating in any study activity, and the study was approved by the OHSU Institutional Review Board (IRB #2765). A total of 76 individuals agreed to participate. Of these, 54 participants searched in Google, Bing, Yahoo, or Ask.com while the software was installed on the computer. However, 12 of these participants did not search during the 3 months before or after completing their neuropsychological evaluation, and thus their Internet search data could not be used. Thus, a total of 42 participants were included in the final analysis. The demographic characteristics of the participants are shown in Table 1. Those participants who completed a search were younger than those who never completed a search, but they were not statistically significantly different in other demographic characteristics. Among those who completed a search, the average age was 81.1 years (SD 10.5), 83% were female (35/42), and 49% had completed college (25/42). The average cognitive z-score (defined in detail under *Cognitive Function* below) in this cohort was 0.16 (SD 0.56), and one participant had a CDR score of at least 0.5, suggesting mild cognitive impairment.

Data and Measures

For each participant, demographics (including age, sex, and education) were collected at baseline, clinical assessments were performed annually, and (beginning in 2015) data were collected continuously from the participants' personal computers using WorkTime. Below is a more detailed description of each of these types of data.

Internet Searches

The WorkTime software installed on participants' computers records the websites visited, the applications used, and the search terms entered when performing Internet-based searches. WorkTime collects data from any search browser on any website such as Target.com and Facebook.com. However, searches on websites other than major search engines do not represent the same type of search query as those in the major search engines. We therefore limited the final dataset to consist only of searches entered in major search engines. A google search revealed that the top search engines are Google, Bing, and Yahoo. We therefore included these three search engines in the dataset. We then reviewed each participant to determine if they frequently conducted searches in other major search engines and found that several participants also conducted searches using Ask.com. We therefore also included searches arising from this search engine in our final dataset.

The final dataset represents an average of 370 days of continuous computer use data (min: 7 days; max: 796 days) from 76 participants. From the time the computer software was installed until the data was pulled for analysis, 54 subjects completed 8565 searches in Bing, Google, Yahoo, and Ask.com, whereas 22 participants with WorkTime installed on their personal computer never completed a search in one of these top search engines during the monitoring period, although some of these participants did conduct searches outside the monitoring period. Because we limited the analysis to the 3 months of Internet search data before or after the in-person cognitive evaluation, only 42 participants who conducted an Internet search during this time were included in the final analyses.

Before analysis, all search terms were cleaned using a 3-step process. First, all unreadable characters were removed from the string of search terms. Such characters include symbols and non-ASCII characters which could not be read or interpreted using standard English text analysis techniques. Where applicable, accented or non-ASCII characters were converted into the ASCII equivalent of the character (eg, "ű" was converted to "u").

Table 1. Demographic characteristics of the participants at baseline.

Characteristics	Participants who completed a criterion search (n=42)	Participants who never completed a search (n=32)	*P* value
Age in years, mean (SD)	81.1 (10.5)	88.9 (6.1)	<.001
Sex, female, n (%)	35 (83)	22 (68.8)	.14
Education in years, mean (SD)	15.5 (2.0)	15.3 (2.5)	.62
Cumulative Illness Rating Scale (CIRS) score, mean (SD)	20.3 (2.6)	20.8 (2.6)	.37
Mini-Mental State Examination (MMSE) score, mean (SD)	29 (1.3)	28.6 (1.7)	.46
Clinical Dementia Rating (CDR) score ≥0.5, n (%)	1 (3)	4 (10.3)	.21
Cognitive z-score, mean (SD)	0.16 (0.56)	0.08 (0.76)	.60

Once these characters were removed, punctuation was also removed, all letters were changed to lowercase, and the complete search was divided into individual, unigram search terms. For example, for the search “Where are cooking classes in Portland?” we first removed the question mark from the end and changed the capital letters in “Where” and “Portland” to lowercase letters. We then divided the entire search into its individual terms: “where,” “are,” “cooking,” “classes,” “in,” and “portland.” After each search was divided into its individual terms, we ran a basic spell checker on each term. Spell checking is necessary as we are interested in understanding how cognitive function relates to variables such as the number of unique terms searched for—a variable that would become over inflated for any participant who regularly misspelled words. This is especially important as spelling may decline with deteriorating cognitive function, potentially masking the true relationship to the generation of the search terms themselves regardless of their correct spelling. To determine whether a word was misspelled, we first tested whether the word appeared in a large corpus of words. If the word did not appear there, we assumed the word was misspelled and endeavored to find the correct spelling. This was done algorithmically by first removing individual letters from the word and testing whether the new word appeared in the large corpus of words. If that did not find a suitable match, we swapped letters that were next to each other in order and tested each newly generated word to see whether it was in the large corpus of words. If we still did not find a match, we added individual letters to the word and tested each new generated word against the corpus of words. If none of these methods found a match, we left the word as is.

The final step in the search term preparation was to stem each word to the root word. For the search example above, this would entail removing the “es” from classes (ie, “classes” becomes “class”) and removing the “ing” from “cooking.” To stem the words, we employed the WordNet lemmatizer that is freely available as part of the Natural Language Toolkit for Python. Unlike other stemming tools, this stemmer first checks WordNet’s expansive dictionary to ensure the stemmed word is an actual word before stemming the suffix.

We calculated three metrics using the cleaned search terms. First, we calculated the average number of unique terms per search, which is defined as the total number of words a participant searched for divided by the total number of searches they performed. Second, we calculated the average length of words searched for as the total number of letters in all words divided by the total number of words searched. These two metrics were computed on a per-subject basis (eg, only data from one subject was used to compute that subject’s average number of unique terms or average word length). Finally, we calculated the average obscurity of the words searched. To compute the average term obscurity, we first calculated the frequency of each searched word across all subjects. Using this frequency, we then calculated the obscurity of each term as the inverse of the frequency. Finally, we calculated each participant’s average term obscurity as the average obscurity score of all words that participant searched for.

Cognitive Function

Standardized, detailed clinical data were collected at baseline for all participants and then annually to appropriately characterize the participants [10]. These clinical data cover four important domains: health status, physical function, cognition, and mood. Global cognitive status was assessed using a composite score including z-scores tabulated from two or three representative neuropsychological tests in each of five cognitive domains. Cognitive domains that were assessed include *working memory*: Letter-Number Sequencing (WMS-III) [29] and Digit Span Backward length (WAIS-R) [30]; *attention/processing speed*: Digit Span Forward length (WAIS-R), Digit Symbol (WAIS-R), and Trail Making Test Part A [31]; *memory*: WMS-R Logical Memory II Story A, WMS-R Visual Reproduction II, and CERAD Word-List Recall [32]; *executive function*: letter fluency (CFL), Trail Making Test Part B [31], and Stroop color-word conflict [33]; and *visual perception/construction*: WAIS-R Block Design, WAIS-R Picture Completion, and WMS-R Visual Reproduction I. Individual test z-scores were calculated using group mean and standard deviations of the raw test scores from all cognitively intact participants at study entry into the ORCATECH cohorts. All individual participant scores were z-normalized, summed, and averaged to obtain a composite score. This latter score represents global cognitive function, hereafter referred to as the cognitive z-score.

Covariates

We included several variables in all models that might confound the relationship between Internet-based search terms and cognitive function. These included age, sex, and years of education, as these are the variables that may relate to cognitive decline.

Data Analyses

We first computed descriptive statistics for all variables. These included the average number of searches conducted on individuals, the average number of words per search, the average number of letters in each word searched for, and the average obscurity of the terms searched.

Next, we ran 3 linear regressions, each with cognitive z-score as the outcome variable. The first linear regression modeled the relationship between cognitive function and number of unique search terms entered per search. The second regression modeled the relationship between the average length of words searched for and cognitive function. The third regression modeled the relationship between cognitive function and average term obscurity. To ensure coefficient estimates were not biased by multicollinearity, the variance inflation factor (VIF), a standard diagnostic tool for assessing the level of collinearity of the independent variables, was computed for all independent variables. The VIF for all variables was below 2.5, indicating any bias from multicollinearity can reasonably be ignored [34]. All regressions controlled for age, sex, and years of education were performed in Stata 13 (Stata Corp) using the function “reg.”

Results

Descriptive Statistics

Participants conducted a median of 22 (interquartile range, IQR 7.3) searches over the 6 month monitoring period (Table 2). The most common terms searched for were “Portland” (n=318 searches), “Oregon” (n=175 searches), “how” (n=40 searches), and “email” (n=36 searches). The mean number of words per search was 3.08 (SD 1.6), and the longest search contained 22 words. Across all words searched for, participants averaged 5.77 (SD 2.2) letters per word. The average term obscurity across participants was 0.25 (SD 0.1).

The richness of the search term dataset is demonstrated in Figure 1 where searches are represented in a social network graph. The figure was created in Gephi 0.8.2, an open source software designed for visualizing social network diagrams. In this figure, each unique term a participant searched for is represented by an individual node, and nodes are joined together if they appeared in the same search (thicker edges indicate they appeared together more frequently). Nodes were sized by their degree such that larger nodes had more unique connections and were colored to represent “communities,” where the communities were determined using Gephi’s modularity function with a resolution of 1.0. The terms ‘Portland” and “Oregon” were searched for so frequently that they overshadowed the rest of the terms and were therefore removed from the graph to allow better visualization of the rest of the network. From the graph, it is clear that people frequently search for “photo” and ”how” which may indicate that people are using the Internet to see pictures of things and to determine how to do things.

Table 2. Descriptive statistics of the variables included in the model.

Variables	Statistic	Range (min-max)
Number of searches, median (IQR, interquartile range)	22 (7.3)	(1-718)
Words per search, mean (SD)	3.08 (1.57)	(1-22)
Letters per word, mean (SD)	5.77 (2.23)	(1-28)
Word obscurity, mean (SD)	0.25 (0.11)	(0.52-0.04)

Figure 1. A social network diagram of participant searches over the past year. Search terms are connected to each other if they appeared in the same search, and stronger connections indicate they appeared more frequently together. Each term is sized by the degree of the node, which represents the number of unique terms that are connected to that term. Terms are colored by community, where terms that are frequently searched for together are grouped into the same community.

Linear Regression Analyses

The results of all three linear regressions are presented in Table 3. Note that in this table, the beta coefficients represent the amount cognitive z-score will change for a unit change in each independent variable. For all outcome variables, we present one-sided *P* values commensurate with the directional relationships hypothesized before running the models.

In the first model, we tested whether the higher cognitive function was associated with more unique search terms entered per search. As shown in Table 3, our results supported this hypothesis: for each additional unique word searched for per search, participants score 0.39 points higher on their cognitive z-score (*P*=.002). To put this in perspective, with this beta coefficient, the model would predict that the participant who averaged the most unique terms per search of 3.1 terms would score 1.01 points higher on their cognitive z-score as compared to the individual who averaged the least unique terms per search of 0.5 terms. This difference is significant considering the range of cognitive z-scores is from −1.15 to 1.2. The R^2 for this model was 0.46. A scatter plot of the relationship between cognitive z-score and the average number of unique words per search can be visualized in Figure 2.

In our second model, we tested whether the higher cognitive function was associated with using longer words when searching. As shown in Table 3, our results did not support this hypothesis (*P*=.21). Although direction of the coefficient was in the hypothesized direction, the relationship was not significant in this cohort. The R^2 for this model was 0.25. A scatter plot of the relationship between cognitive z-score and the average number of letters per word can be visualized in Figure 2.

In our final regression, we tested whether individuals with a higher cognitive function would use more obscure words when they searched the Internet. As shown in Table 3, our results supported this hypothesis: for each additional unit increase in the average obscurity of the words searched for, participants scored 1.39 points higher on their cognitive battery (*P*=.02). To put this in perspective, with this beta coefficient, the model would predict that the participant with the highest average term obscurity of 0.52 would score 0.66 points higher on their cognitive z-score compared to the participant with the lowest average term obscurity of 0.044. The R^2 for this model was 0.32. A scatter plot of the relationship between cognitive z-score and the average term obscurity can be visualized in Figure 2.

In all models, age was significantly related to cognitive function such that older individuals had a lower cognitive function. Sex and education were not significantly associated with cognitive function in any model.

Table 3. Results of the three linear regressions relating Internet searches to cognitive function.

Characteristics	Model 1	Model 2	Model 3
	Beta coefficient (SD)	Beta coefficient (SD)	Beta coefficient (SD)
Constant	.75 (0.96)	1.24 (1.10)	1.53 (0.98)
Age	−.024 (0.007)[a]	−.024 (0.008)[b]	−.024 (0.007)[c]
Sex (Female)	.27 (0.20)	.19 (0.23)	.136 (0.22)
Education	.016 (0.038)	.006 (0.043)	.005 (0.041)
Number of Unique Terms per Search	.39 (0.13)[b]		
Average Number of Letters per Word		.084 (0.806)	
Average Term Obscurity			1.39 (0.68)[d]

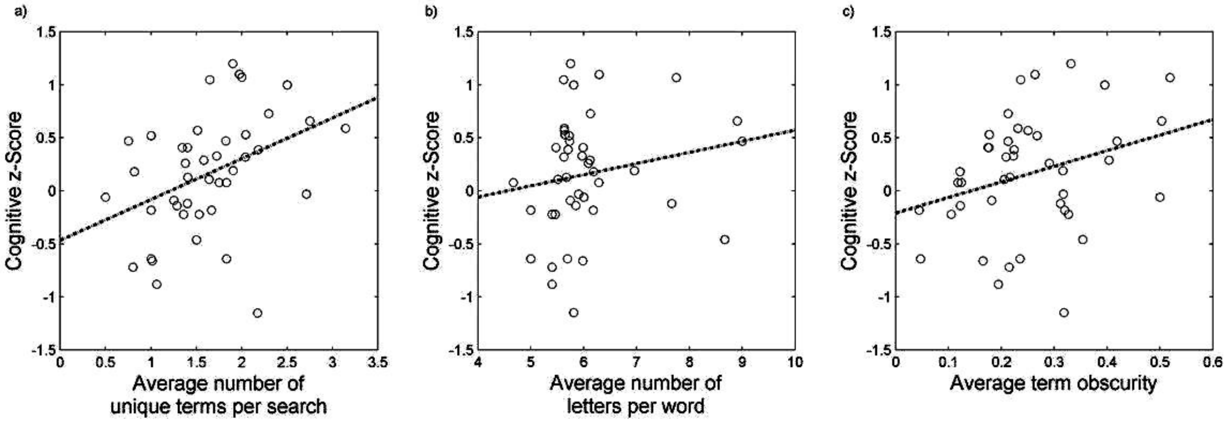
[a] *P*=.001.

[b] *P*=.004.

[c] *P*=.002.

[d] *P*=.02.

Figure 2. Scatter plots of the relationships between cognitive function and (a) average number of unique terms per search, (b) the average number of letters per word, and (c) the average term obscurity. The observed regression line for each relationship is also plotted as a dashed line.



Discussion

Study Overview

In this study, we used WorkTime to continuously monitor computer use in a sample of 74 older adults. WorkTime records the terms people search for whenever they conduct an Internet-based search. Using the search terms data from 42 subjects who completed at least one search during the 6 months surrounding (3 months before or 3 months after) a cognitive evaluation, we demonstrated cognitive function is tied to both the average number of unique terms entered per search and obscurity of the searched words. These results present the first time to our knowledge that a continuous aspect of language use, Internet search terms, has been related to the cognitive abilities of older adults. A follow-up analysis should assess this relationship longitudinally in presymptomatic older adults to determine whether continuous assessment of Internet search terms can be used to identify individuals who will eventually transition to mild cognitive impairment.

Principal Findings

Our first hypothesis was that individuals with a superior cognitive function would employ more unique terms in their searches. Our results supported this hypothesis, the results of which might be considered consistent with the lexical or generative fluency needed to perform well in the standard psychometric tasks of category and phonemic fluency. Previous studies have established that phonemic and category fluency test performance decline with a change in cognitive function [18,19] including transitions leading to Mild Cognitive Impairment [35]. Category fluency represents the ability to name members of a category (eg, animals), whereas phonemic fluency represents the ability to name words that begin with a certain letter. Both are typically assessed by having participants name as many objects or words as possible in a fixed period of time. Due to the close relationship between cognitive decline and language use, these tests are often part of standard batteries of cognitive tests designed to detect MCI or Alzheimer disease. One may consider the task of generating search terms to draw upon similar cognitive resources, and thus the analysis of search terms generated over time and presented here represents the first time this measure of fluency (fluency of Internet search terms or “FIST”) has been linked with cognitive function.

Our second hypothesis was that individuals with a superior cognitive function would search using longer words. However, our results did not support this hypothesis. This is also consistent with previous studies that have found that word length is not as closely related to picture naming ability as term obscurity [18]. That is, vocabulary in Alzheimer disease may decline disproportionately with word obscurity or word familiarity rather than word complexity. Indeed, our final hypothesis was that individuals with a superior cognitive function would search using more obscure search terms. Our results supported this hypothesis independent of years of formal education, which is consistent with previous work linking the decline in vocabulary to cognitive function [23].

Limitations

This study has several limitations. Of note, the participants included in this study were primarily white, well-educated, and relatively healthy older adults. The results reported here may not generalize to other populations. The sample size was also small, therefore we controlled for only a small number of variables. Future studies should investigate whether variables such as social network size, social economic status, computer fluency, or medication use have any effect on the results reported here. Computer fluency may be especially important as there could be significant differences in computer use not due to cognitive decline but due to familiarity and exposure to the computer [36].

In addition, we limited the search terms included in the model to only those arising from four major search engines: Google, Bing, Yahoo, and Ask.com. Although these are the primary search engines used by these older adults, it is possible that not all searches were captured as some participants may search the Internet using other, less common search engines.

We also had no way to determine whether participants used an autofill to auto complete their searches. Several search browsers provide the option to give suggestions on the potential remaining terms in a search, typically using popular search queries from both the participants and the greater public to inform the suggestions. If participants were using such software when performing searches, it would inject artificial noise into the search terms dataset. Future studies should verify the results presented here in a dataset where the auto complete function was disabled in all participant browsers.

We employed a basic stemmer and spell checker. These utilities ensured that conjugates of words (for example “running” is a conjugate of “run”) would not be counted twice in the term frequency dictionary, and that misspelled words would not be treated as highly obscure terms when they are actually very frequent but misspelled. However, neither of these utilities performed perfectly. For example, while the stemmer correctly stemmed “wolves” to “wolf,” it incorrectly stems “dies” to “dy.” A more sophisticated spell checker and stemmer may enhance future studies.

Our measure of term obscurity was simply the inverse of the frequency with which the word appeared in the search dataset. This was necessary as multiple words such as “Gmail” are common on the Internet but not characterized in common measures of word frequency or rarity. However, because the subject searches were used both to develop the term frequency dictionary and assess the average word obscurity for each subject, it is possible that individuals who searched the Internet more had a lower average term obscurity as their search phrases and terms were entered more frequently into the dataset. Indeed, the number of searches was negatively correlated with the average word obscurity ($r=-0.23$). However, follow-up analysis revealed that the number of searches in the dataset was not related to the cognitive z-score of the individual ($P=.97$). Nevertheless, future studies should normalize the word frequency per subject to compute the relative obscurity of each word.

WorkTime can monitor not only the terms people search for on the Internet but also detailed aspects of computer use such as the time spent in online games or in social websites. Thus, future studies may benefit from assessing the relationship between cognitive function and multiple aspects of computer use, especially as recent studies have demonstrated that the total time spent on the computer and the number of computer sessions is related to cognitive decline [13,37,38]. Variability in computer use has also been linked to cognitive decline [13], but few, if any, studies have assessed more detailed aspects of computer use (eg, total time in online games) and their relationship to cognitive function, especially using an objective monitoring software.

Conclusions

This work uniquely assessed the relationship between everyday language function as demonstrated through Internet based searches and cognitive function. Several studies have shown a close relationship between language abilities such as vocabulary size, verbal fluency, and semantic ability and overall change in cognitive function. These functions are typically assessed through standardized episodically administered cognitive tests. Previous studies have not been able to assay such cognitive constructs at a level that provides a scalable early detection approach for cognitive decline. This is likely in part due to the lack of methods providing frequent everyday samples of language use. The approach proposed here takes advantage of computer software that makes it possible to continually and unobtrusively capture aspects of language and related complex cognitive activity during routine computer use. In addition, unlike prior work, the data to be analyzed is inherently ecologically valid as it is the individuals' everyday function that is being assessed. By building on the relationships demonstrated here, it may be possible to develop a system that detects the prodromal stage of Alzheimer disease by continuously monitoring the terms people search for on the Internet along with other aspects of everyday computer use [7,13,14]. This could be accomplished with algorithms that run routinely and securely in the background similar to virus detection software. By developing a naturalistic technique to assess the earliest symptoms of cognitive change, this approach has the potential to significantly advance the diagnostic and assessment process and provide a novel mechanism that can be used in improving the conduct of clinical trials and care for the development of AD treatments. This would have significant and far-reaching effects on older adults experiencing cognitive decline, their families, and the health care system as a whole.

Acknowledgments

This study was supported, in part, by Oregon citizens through the Alzheimer's Disease Research Fund of the Oregon Charitable Checkoff Program. This fund is administered by the Oregon Partnership for Alzheimer's Research. In addition, National Institute on Aging grants R01 AG042191, P30 AG008017, and P30 AG024978 contributed significantly in supporting this research.

Conflicts of Interest

None declared.

References

1. Alzheimer's Association. 2014 Alzheimer's disease facts and figures. Alzheimers Dement 2014 Mar;10(2):e47-e92. [Medline: 24818261]
2. Cummings JL, Morstorf T, Zhong K. Alzheimer's disease drug-development pipeline: few candidates, frequent failures. Alzheimers Res Ther 2014 Jul 03;6(4):37 [FREE Full text] [doi: 10.1186/alzrt269] [Medline: 25024750]
3. Aisen PS, Vellas B, Hampel H. Moving towards early clinical trials for amyloid-targeted therapy in Alzheimer's disease. Nat Rev Drug Discov 2013 Apr;12(4):324. [doi: 10.1038/nrd3842-c1] [Medline: 23493086]
4. Deak F, Freeman WM, Ungvari Z, Csiszar A, Sonntag WE. Recent developments in understanding brain aging: implications for Alzheimer's disease and vascular cognitive impairment. J Gerontol A Biol Sci Med Sci 2016 Jan;71(1):13-20 [FREE Full text] [doi: 10.1093/gerona/glv206] [Medline: 26590911]
5. Serra L, Musicco M, Cercignani M, Torso M, Spanò B, Mastropasqua C, et al. Cognitive reserve and the risk for Alzheimer's disease: a longitudinal study. Neurobiol Aging 2015 Feb;36(2):592-600. [doi: 10.1016/j.neurobiolaging.2014.10.010] [Medline: 25433459]
6. Dodge HH, Zhu J, Mattek NC, Austin D, Kornfeld J, Kaye JA. Use of high-frequency in-home monitoring data may reduce sample sizes needed in clinical trials. PLoS One 2015 Sep 17;10(9):e0138095. [doi: 10.1371/journal.pone.0138095] [Medline: 26379170]
7. Seelye A, Hagler S, Mattek N, Howieson DB, Wild K, Dodge HH, et al. Computer mouse movement patterns: a potential marker of mild cognitive impairment. Alzheimers Dement (Amst) 2015 Dec 01;1(4):472-480 [FREE Full text] [doi: 10.1016/j.dadm.2015.09.006] [Medline: 26878035]
8. Petersen J, Austin D, Mattek N, Kaye J. Time out-of-home and cognitive, physical, and emotional wellbeing of older adults: a longitudinal mixed effects model. PLoS One 2015 Oct 05;10(10):e0139643 [FREE Full text] [doi: 10.1371/journal.pone.0139643] [Medline: 26437228]
9. Lyons BE, Austin D, Seelye A, Petersen J, Yeargers J, Riley T, et al. Pervasive computing technologies to continuously assess Alzheimer's disease progression and intervention efficacy. Front Aging Neurosci 2015 Jun 10;7:102 [FREE Full text] [doi: 10.3389/fnagi.2015.00102] [Medline: 26113819]
10. Kaye JA, Maxwell SA, Mattek N, Hayes TL, Dodge H, Pavel M, et al. Intelligent systems for assessing aging changes: home-based, unobtrusive, and continuous assessment of aging. J Gerontol B Psychol Sci Soc Sci 2011 Jul;66(Suppl 1):i180-i190 [FREE Full text] [doi: 10.1093/geronb/gbq095] [Medline: 21743050]
11. Zickuhr K, Madden M. Pew Internet. 2012 Jun 06. Older adults and internet use URL: http://www.pewinternet.org/2012/06/06/older-adults-and-internet-use/ [accessed 2017-08-09] [WebCite Cache ID 6sbeNFK8e]
12. Smith A. Pew Internet. 2014 Apr 03. Older adults and technology use URL: http://www.pewinternet.org/2014/04/03/older-adults-and-technology-use/ [accessed 2017-08-10] [WebCite Cache ID 6sbecowMw]

13. Kaye J, Mattek N, Dodge HH, Campbell I, Hayes T, Austin D, et al. Unobtrusive measurement of daily computer use to detect mild cognitive impairment. Alzheimers Dement 2014 Jan;10(1):10-17 [FREE Full text] [doi: 10.1016/j.jalz.2013.01.011] [Medline: 23688576]
14. Austin D, Jimison H, Hayes T, Mattek N, Kaye J, Pavel M. Measuring motor speed through typing: a surrogate for the finger tapping test. Behav Res Methods 2011 Dec;43(4):903-909 [FREE Full text] [doi: 10.3758/s13428-011-0100-1] [Medline: 21494919]
15. Austin D, Petersen J, Jimison H, Pavel M. A state-space model for finger tapping with applications to cognitive inference. In: Conf Proc IEEE Eng Med Biol Soc. 2012 Presented at: 34th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC); 28 Aug - 01 Sep 2012; San Diego, CA, USA p. 21-24. [doi: 10.1109/EMBC.2012.6345861]
16. Gomez RG, White DA. Using verbal fluency to detect very mild dementia of the Alzheimer type. Arch Clin Neuropsychol 2006 Dec;21(8):771-775 [FREE Full text] [doi: 10.1016/j.acn.2006.06.012] [Medline: 17011743]
17. Martin A, Fedio P. Word production and comprehension in Alzheimer's disease: the breakdown of semantic knowledge. Brain Lang 1983 May;19(1):124-141. [Medline: 6860932]
18. Beber BC, Cruz AN, Chaves ML. A behavioral study of the nature of verb production deficits in Alzheimer's disease. Brain Lang 2015 Oct;149:128-134. [doi: 10.1016/j.bandl.2015.07.010] [Medline: 26291288]
19. Mueller KD, Koscik RL, LaRue A, Clark LR, Hermann B, Johnson SC, et al. Verbal fluency and early memory decline: results from the Wisconsin registry for Alzheimer's prevention. Arch Clin Neuropsychol 2015 Aug;30(5):448-457 [FREE Full text] [doi: 10.1093/arclin/acv030] [Medline: 26025231]
20. Kemper S, LaBarge E, Ferraro FR, Cheung H, Cheung H, Storandt M. On the preservation of syntax in Alzheimer's disease. Evidence from written sentences. Arch Neurol 1993 Jan;50(1):81-86. [Medline: 8418805]
21. Garrard P, Maloney LM, Hodges JR, Patterson K. The effects of very early Alzheimer's disease on the characteristics of writing by a renowned author. Brain 2005 Feb;128(Pt 2):250-260. [doi: 10.1093/brain/awh341] [Medline: 15574466]
22. Laws KR, Duncan A, Gale TM. 'Normal' semantic-phonemic fluency discrepancy in Alzheimer's disease? A meta-analytic study. Cortex 2010 May;46(5):595-601. [doi: 10.1016/j.cortex.2009.04.009] [Medline: 19560132]
23. Fraser KC, Meltzer JA, Rudzicz F. Linguistic features identify Alzheimer's disease in narrative speech. J Alzheimers Dis 2016;49(2):407-422. [doi: 10.3233/JAD-150520] [Medline: 26484921]
24. Szatloczki G, Hoffmann I, Vincze V, Kalman J, Pakaski M. Speaking in Alzheimer's disease, is that an early sign? Importance of changes in language abilities in Alzheimer's disease. Front Aging Neurosci 2015 Oct 20;7:195 [FREE Full text] [doi: 10.3389/fnagi.2015.00195] [Medline: 26539107]
25. Roark B, Mitchell M, Hosom JP, Hollingshead K, Kaye J. Spoken language derived measures for detecting mild cognitive impairment. IEEE Trans Audio Speech Lang Process 2011 Sep 01;19(7):2081-2090 [FREE Full text] [doi: 10.1109/TASL.2011.2112351] [Medline: 22199464]
26. Robertson-Lang L, Major S, Hemming H. An exploration of search patterns and credibility issues among older adults seeking online health information. Can J Aging 2011 Dec;30(4):631-645. [doi: 10.1017/S071498081100050X] [Medline: 22085455]
27. Hardt JH, Hollis-Sawyer L. Older adults seeking healthcare information on the Internet. Educ Gerontol 2007 Jun 11;33(7):561-572. [doi: 10.1080/03601270701364628]
28. Flynn KE, Smith MA, Freese J. When do older adults turn to the internet for health information? Findings from the Wisconsin Longitudinal Study. J Gen Intern Med 2006 Dec;21(12):1295-1301 [FREE Full text] [doi: 10.1111/j.1525-1497.2006.00622.x] [Medline: 16995892]
29. Wechsler D. Wechsler memory scale (WMS-III). San Antonio, TX: Psychological Corporation; 1997.
30. Wechsler D. WAIS-R manual: Wechsler adult intelligence scale-revised. San Antonio, TX: Psychological Corporation; 1981.
31. Armitage S. An analysis of certain psychological tests used for the evaluation of brain injury. In: Psychological Monographs. Washington DC: American Psychological Association; 1946.
32. Rosen WG, Mohs RC, Davis KL. A new rating scale for Alzheimer's disease. Am J Psychiatry 1984 Nov;141(11):1356-1364. [doi: 10.1176/ajp.141.11.1356] [Medline: 6496779]
33. Jensen AR, Rohwer Jr WD. The Stroop color-word test: a review. Acta Psychol (Amst) 1966;25(1):36-93. [Medline: 5328883]
34. Allison P. Statistical Horizons. 2012 Sep 10. When Can You Safely Ignore Multicollinearity? URL: http://www.statisticalhorizons.com/multicollinearity [accessed 2017-08-10] [WebCite Cache ID 6sbhEYtxu]
35. Howieson DB, Carlson NE, Moore MM, Wasserman D, Abendroth CD, Payne-Murphy J, et al. Trajectory of mild cognitive impairment onset. J Int Neuropsychol Soc 2008 Mar;14(2):192-198. [doi: 10.1017/S1355617708080375] [Medline: 18282317]
36. van Boekel LC, Peek ST, Luijkx KG. Diversity in older adults' use of the Internet: identifying subgroups through latent class analysis. J Med Internet Res 2017 May 24;19(5):e180 [FREE Full text] [doi: 10.2196/jmir.6853] [Medline: 28539302]
37. Tun PA, Lachman ME. The association between computer use and cognition across adulthood: use it so you won't lose it? Psychol Aging 2010 Sep;25(3):560-568 [FREE Full text] [doi: 10.1037/a0019543] [Medline: 20677884]

38. Slegers K, van Boxtel MPJ, Jolles J. Computer use in older adults: determinants and the relationship with cognitive change over a 6 year episode. Comput Human Behav 2012 Jan;28(1):1-10. [doi: 10.1016/j.chb.2011.08.003]
39. Gephi. Gephi : An Open Source Software for Exploring and Manipulating Networks URL: https://gephi.org/publications/gephi-bastian-feb09.pdf [accessed 2017-08-29] [WebCite Cache ID 6t4jsNhSV]

Abbreviations

AD: Alzheimer disease
CDR: Clinical Dementia Rating
CIRS: Cumulative Illness Rating Scale
MCI: Mild Cognitive Impairment
MMSE: Mini-Mental State Examination
VIF: Variance Inflation Factor

Edited by G Eysenbach; submitted 16.03.17; peer-reviewed by K Fraser, S Bonasera, Q Zhang; comments to author 27.04.17; revised version received 21.06.17; accepted 29.07.17; published 06.09.17.

Please cite as:
Austin J, Hollingshead K, Kaye J
Internet Searches and Their Relationship to Cognitive Function in Older Adults: Cross-Sectional Analysis
J Med Internet Res 2017;19(9):e307
URL: http://www.jmir.org/2017/9/e307/
doi:10.2196/jmir.7671
PMID:28877864

Original Paper

Self-Monitoring Kidney Function Post Transplantation: Reliability of Patient-Reported Data

Céline van Lint[1], MSc; Wenxin Wang[2], MSc; Sandra van Dijk[3], MSc, PhD; Willem-Paul Brinkman[2], MSc, PhD; Ton JM Rövekamp[4], BSc; Mark A Neerincx[2,4], PhD; Ton J Rabelink[1], MD; Paul JM van der Boog[1], MD, PhD

[1]Leiden University Medical Center, Department of Nephrology, Leiden, Netherlands
[2]Interactive Intelligence Group, Delft University of Technology, Delft, Netherlands
[3]Department of Health, Medical and Neuropsychology and Behavioural Sciences, Leiden University, Leiden, Netherlands
[4]Dutch Organization for Applied Scientific Research (TNO), the Hague, Netherlands

Corresponding Author:
Céline van Lint, MSc
Leiden University Medical Center
Department of Nephrology
Albinusdreef 2
Leiden, 2333 ZA
Netherlands
Phone: 31 715262148
Fax: 31 715266868
Email: c.l.van_lint@lumc.nl

Abstract

Background: The high frequency of outpatient visits after kidney transplantation is burdensome to both the recovering patient and health care capacity. Self-monitoring kidney function offers a promising strategy to reduce the number of these outpatient visits.

Objective: The objective of this study was to investigate whether it is safe to rely on patients' self-measurements of creatinine and blood pressure, using data from a self-management randomized controlled trial.

Methods: For self-monitoring creatinine, each participant received a StatSensor Xpress-i Creatinine Meter and related test material. For self-monitoring blood pressure, each participant received a Microlife WatchBP Home, an oscillometric device for blood pressure self-measurement on the upper arm. Both devices had a memory function and the option to download stored values to a computer. During the first year post transplantation, 54 patients registered their self-measured creatinine values in a Web-based Self-Management Support System (SMSS) which provided automatic feedback on the registered values (eg, seek contact with hospital). Values registered in the SMSS were compared with those logged automatically in the creatinine device to study reliability of registered data. Adherence to measurement frequency was determined by comparing the number of requested with the number of performed measurements. To study adherence to provided feedback, SMSS-logged feedback and information from the electronic hospital files were analyzed.

Results: Level of adherence was highest during months 2-4 post transplantation with over 90% (42/47) of patients performing at least 75% of the requested measurements. Overall, 87.00% (3448/3963) of all registered creatinine values were entered correctly, although values were often registered several days later. If (the number of) measured and registered values deviated, the mean of registered creatinine values was significantly lower than what was measured, suggesting active selection of lower creatinine values. Adherence to SMSS feedback ranged from 53% (14/24) to 85% (33/39), depending on the specific feedback.

Conclusions: Patients' tendency to postpone registration and to select lower creatinine values for registration and the suboptimal adherence to the feedback provided by the SMSS might challenge safety. This should be well considered when designing self-monitoring care systems, for example by ensuring that self-measured data are transferred automatically to an SMSS.

(J Med Internet Res 2017;19(9):e316) doi:10.2196/jmir.7542

KEYWORDS

self-care; kidney transplantation; creatinine; patient compliance; data accuracy; patient reported outcomes

Introduction

After kidney transplantation, an early detection of transplant failure is mandatory to minimize permanent damage to the transplanted organ. For kidneys, blood level of creatinine is considered the most important indicator of kidney function [1]. Patients therefore have their serum creatinine checked on average 20 times during the first year post transplantation. As hypertension is both a potential indicator of decreased kidney function and an important risk factor for kidney graft failure [2-5], blood pressure needs extensive monitoring too. If patients were enabled to monitor both parameters at home, this would have important advantages. Self-monitoring could improve the speed of rejection detection as measurements can take place more frequently while at the same time the high number of outpatient visits could be reduced and replaced by telephonic consults. Furthermore, giving patients a more active role in their own care through self-monitoring has been shown to be of clinical benefit for a wide range of patients with chronic disease [6-15] and to lead to a higher quality of life [16-19]and more patient empowerment [7,19-22].

A pilot study of our own group showed that self-monitoring of both blood pressure and creatinine is very well accepted among patients, suggesting that at-home monitoring after transplantation offers a promising strategy [23]. For self-monitoring to be a safe alternative to regular face-to-face follow-up, however, patients need to adhere to a monitoring schedule, report test results accurately, and act upon test results if these suggest graft failure may occur. This is important for all patients who engage in self-monitoring, but especially for patients who are transplanted. As most patients who develop graft rejection are asymptomatic and present with an increased serum creatinine only, frequent measuring is essential to make the difference between treatment in time and damage to or even loss of the kidney transplant. Level of adherence to a self-monitoring schedule has been shown to vary widely in other disease populations [24-28]. Further, for self-measured values to be clinically useful, they need to be reported accurately. Several studies in different study populations have shown that caution is warranted when using patient-reported data for making clinical decisions as a considerable number of patients report values that do not sufficiently represent their actual measurements [29-34].

To the best of our knowledge, no studies have assessed the reliability and accuracy of patient-generated creatinine data or looked at the level of adherence to a protocol of self-monitoring creatinine. This is unfortunate, as the introduction of self-monitoring offers a good opportunity to improve post-transplantation care. Our first research goal was to investigate the level of adherence of kidney transplant patients to a creatinine monitoring schedule. Our second research goal was to determine the reliability of the creatinine values that were registered in a Web-based self-management support system (SMSS). As this SMSS automatically provided instructions for further actions (eg, continue regular schedule or contact the hospital) upon registration of a new creatinine value, our final research goal was to determine whether patients adhered to the system’s instructions.

Methods

Patients and Study Design

The data used in this study were obtained from the ADMIRE project (Assessment of a Disease management system with Medical devices In REnal disease), a cooperation between the Leiden University Medical Centre (LUMC), the Technical University of Delft, and the Dutch Organization for Applied Scientific Research (registered in the Dutch Trial Register: NTR3548). This extensive project comprised the technical development of an SMSS in which several studies were performed to optimize the system to suit patients’ needs and wishes, as well as a prospective randomized controlled trial (RCT) to study whether self-monitoring kidney function supported by an SMSS can replace part of regular outpatient care without compromising on the quality of care. The study protocol was approved by the Medical Ethics Committee of the LUMC.

Patients were eligible for participation in the RCT if they were about to receive a donor kidney or recently received one, were ≥18 years of age, mastered the Dutch language sufficiently, had access to the Internet, and had a creatinine level of ≤300 µmol/l within 4 weeks post transplantation. Patients were excluded if they were visually impaired or were considered ineligible by their treating physician (eg, due to mental retardation, a history of noncompliance to treatment). We therefore had a selection of patients that seemed most capable for engaging in self-monitoring.

Recruitment of living donor recipients took place during a pretransplant consultation with a nurse-practitioner. Recipients of a postmortem kidney were recruited during their post-transplantation stay in the hospital by the primary investigator (CvL). After this face-to-face introduction, patients received a written explanation of the study with an informed consent form. If a signed informed consent was not returned within 2 weeks from the recruitment date, patients were contacted telephonically to inquire whether they were interested in participating. After signing informed consent, each participant was assigned a study number. Incoming informed consents were treated in consecutive order. Study numbers were allocated to either the intervention or the control group according to a preset randomization schedule which was created by a medical statistician. The randomization procedure was blinded for the project members directly involved in patient recruitment.

For this study, only participants randomized to the intervention group were included.

Intervention

Devices and Self-Management Support System

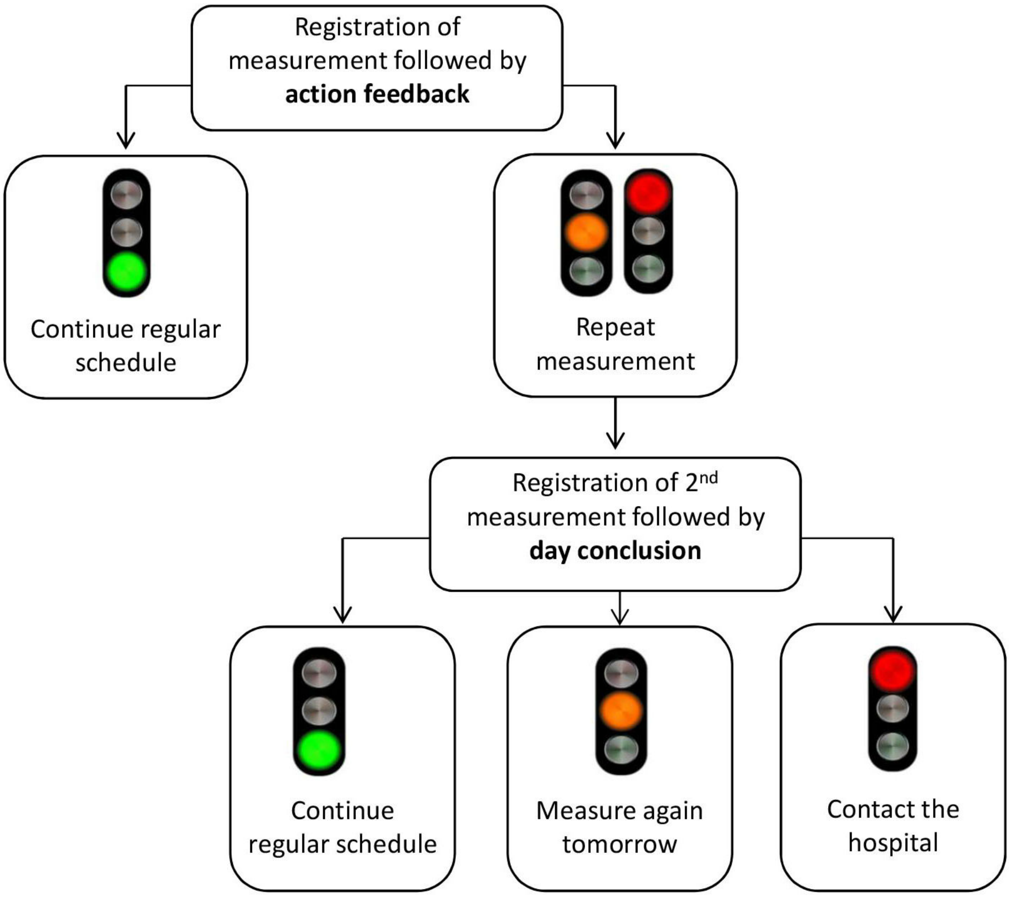
For self-monitoring creatinine, each participant received a StatSensor Xpress-i Creatinine Meter (Nova Biomedical, Waltham, USA) and related test material (ie, test strips, control solution to test the quality of the strips, and safety lancets for capillary blood sampling). On the basis of a drop of blood of 1.2 µL, the StatSensor can show either current level of creatinine or estimated glomerular filtration rate (eGFR). At our medical center, clinicians usually communicate level of creatinine to kidney transplant patients, and the device was set to show creatinine.

For self-monitoring blood pressure, each participant received a Microlife WatchBP Home (Microlife, Heerbrugg, Switzerland), an oscillometric device for blood pressure self-measurement on the upper arm. Both devices had a memory function and the option to download stored values to a computer.

A Web-based SMSS was available for all patients in the intervention group. This SMSS entailed an e-learning module instructing patients on how to use the SMSS system, that is (a) how to perform creatinine measurements at home, (b) how to register self-measured values in the SMSS (both creatinine and blood pressure), and (c) how to respond to messages from the automatic feedback system to support patients' interpretation of the creatinine trends. Figure 1 provides an overview of the feedback process. The feedback appeared directly after registration of a new creatinine value and consisted of a traffic light with corresponding text. Per day, a maximum of two creatinine values could be registered. After registration of the first value of the day, a green light indicated that there was no reason for concern and was associated with the advice to just continue regular measurement frequency. This was termed the day conclusion, as no further actions were required for the concerning day. The appearance of an orange or red light (in case the newly registered value was respectively >15% or >20% higher than mean of the previous five values) directly after registration of the first value of the day indicated that there was some reason for concern. The system's advice was then to perform and register a second measurement to confirm the first measurement. This was termed action feedback, as it required an immediate action. After registration of a second measurement, an appearing green light indicated that there was no further reason for concern. In this case, patients were advised to continue their regular monitoring frequency. Alternatively, an orange light indicated that there was some reason for concern and patients were advised to measure again tomorrow. Finally, a red light indicated that there was reason for concern and patients were advised to contact the hospital. Feedback given after the registration of a second measurement was also termed day conclusion, as no further values could be registered. See Figure 1 for an overview of all possible feedback combinations.

A link was created between the SMSS and the electronic hospital system in order for the registered creatinine values to be visible for the treating nephrologist(s). Nephrologists did not receive a copy of the supplied SMSS feedback.

Figure 1. Possible feedback combinations in the Self-Management Support System (SMSS).



Procedure

The time schedule for providing instructions depended on whether patients participated in the living donor program or whether they received a kidney from a deceased donor. Two weeks before a scheduled transplantation, patients received account details to log in to the SMSS and use the e-learning module to prepare for self-monitoring. Recipients of a kidney from a deceased donor received account details during their post-transplantation stay in the hospital. A laptop was available to use the e-learning modules. After being virtually instructed through e-learning, all patients received the creatinine device and supplementary face-to-face instructions. Special attention was paid to the fact that patients had to take action themselves upon the system's feedback, as their nephrologist(s) would only check the home-based creatinine values in advance of or during an outpatient visit or a telephonic consult. Patients were then encouraged to practice using the creatinine device during the remainder of their hospital stay.

Home-based creatinine measurements had to be performed according to a fixed frequency, being daily during the first 4 weeks (phase 1), every other day for weeks 5-9 (phase 2), twice a week for weeks 10-15 (phase 3), and weekly from week 16 onward (phase 4). This scheme was based upon the usual frequency of laboratory testing, which decreases when time since transplantation increases. However, as the creatinine device tends to be less accurate than laboratory tests [35], the usual frequency of laboratory testing was multiplied with a factor 7 to obtain a more reliable trend. After measuring, patients registered the results in the SMSS.

In addition to conducting regular creatinine measurements, patients were advised to perform a test measurement when opening a new bottle of test strips. These measurements could be termed test measurements by pressing a designated button on the creatinine device.

From week 8 after transplantation on, every other face-to-face outpatient visit with regular hospital-based laboratory measurements was replaced by a telephonic consult to discuss self-monitored creatinine and blood pressure. Although regular face-to-face visits also include other laboratory measurements (eg, trough levels of immunosuppressive medication), these analyses do not need to be performed in the same frequency as for creatinine due to their (expected) little variation over a short period of time. It was therefore deemed unnecessary to replace these other laboratory measurements with a home-based alternative.

To remind nephrologists of scheduling a telephone consult instead of a face-to-face visit, a short note asking for the next appointment to be a telephonic one was shown repetitively in a patient's electronic hospital file. It was, however, up to the treating nephrologist to judge whether a patient's condition allowed for a telephonic consult to take place or whether a face-to-face visit was needed.

At the end of the intervention period of one year, all patients were invited to bring their creatinine device to download logged data. This data included test results, date and time of all performed measurements, and, if applicable, an indication of whether a specific value was termed a test measurement. Further, data that were automatically logged in the SMSS were downloaded including the registered value(s), date of performed measurement (according to the patient), date of registration, and the feedback that was supplied after each newly registered creatinine value.

Measures

Patients completed a questionnaire at baseline to collect demographic characteristics. The read-out data from the creatinine device and the data that were logged in the SMSS were combined using date of measurement. For the creatinine device, measurement date was the date of measurement performance that was registered automatically in the device memory. For the SMSS, measurement date was the date of measurement performance according to the patient.

Statistical Analyses

Adherence to Measurement Frequency

To assess whether patients adhered to the measurement protocol, we separated adherence according to device-logged data (did patients perform the requested number of measurements?) and adherence to SMSS-logged data (did patients register the requested number of measurements in the SMSS?). If applicable, paired *t* tests were conducted to compare means using SPSS 22.0 for Windows (IBM Corp, Armonk, NY). In these cases $P<.05$ was considered statistically significant. For adherence according to device-logged data, we calculated the number of days with measurements per patient per phase and compared this with the number of requested measurement days. Number of requested measurement days was 28 during phase 1 (4 weeks), 15 during phase 2 (5 weeks), 12 during phase 3 (6 weeks), and 37 during phase 4 (37 weeks). To make it easier to interpret the results, level of adherence was divided in four subcategories for this study:

1. Extremely nonadherent: measurements performed during less than 25% of the requested days
2. Nonadherent: measurements performed during 25-74% of the requested measurement days
3. Adherent: measurements performed during 75-100% of the requested measurement days
4. Overadherent: measurements performed more frequently than requested (ie, > 100%).

This same procedure was used to calculate the level of adherence to registration of measurement in the SMSS, that is, whether patients registered measurements on the requested number of days.

Moment of Registration

Date of measurement (derived from device-logged data) was compared with the date of registration of this measurement (derived from SMSS-logged data). Per patient the average number of days delay between measurement and registration was calculated. Furthermore, we investigated whether delayed registration was related to the stability of creatinine level by comparing feedback that was generated by the SMSS in case of registration on day of measurement with feedback that was generated when registration was delayed.

Reliability: Correctness and Representativeness of Registered Data

The reliability of registered data is determined by both the correctness and the representativeness of registered values. To study correctness of the registered data, we investigated the one-on-one correspondence between registered and measured values on a given day. Three different categories were distinguished:

1. Reliable SMSS registrations, in case a value that was registered in the SMSS corresponded to the device logged value on a given date. Only days with an equal number of measurements logged in the device and SMSS were taken into account.
2. Noncorrespondence, in case an SMSS registered value did *not* correspond to the device-logged value on a given date. Only days with an equal number of measurements in the device-logged and SMSS-logged data were selected. All cases of noncorrespondence were carefully checked for potential causes of the deviance (eg, wrong combination of date and measured value, typo, rounding off). The cases where no potential cause was found were termed incorrect entries. For each patient, a mean level of creatinine was calculated for the values that were actually measured and for the values that were registered using cases of incorrect entry only. A paired *t* test was performed to compare these means. Total and median number of noncorresponding values was calculated per patient. Patients with a high number of noncorresponding values were selected for further exploration.
3. Phantom values, in case a value was registered in the SMSS on a given date while according to the data stored in the device no measurement was performed on that specific date. All potential phantom values were thoroughly checked for alternative explanations before it was concluded that there was no relation with values that had been measured by the patient. A paired *t* test was performed using the mean of the phantom values versus the mean of all measured creatinine values per patient.

Furthermore, to get a reliable impression of a creatinine level over time (trend), the SMSS registered values need to represent what was actually measured. It is therefore important to know how often a measured value was not registered in the SMSS and whether the unregistered values differed in any way from the registered values. The measured values not being registered in the SMSS were split into two categories:

1. Omissions, in case one or more measurements were performed on a given date, but no value was registered in the SMSS. Total and median number of omissions per patient was calculated. For each patient, we calculated a mean level of creatinine for the values that were both measured and registered and a mean level of creatinine for the values that were measured on days without any registration. A paired *t* test was performed to compare these means.
2. Selection of measurements, which is the case when the number of performed measurements that is stored in the device is higher than the number of registered measurements on a given date. Therefore only days with a difference between the number of measured and the number of registered creatinine values were selected (eg, three measurements stored in the device and one value registered in the SMSS). We then calculated per patient the mean of all values stored in the device *and* registered in the SMSS and the mean of all values stored in the device, *but not* registered in the SMSS. A paired *t* test was performed to compare these means.

Adherence to Feedback

After registration of a creatinine value in the SMSS, patients received an automatically generated advice on the necessary action to take (see Figure 1). To investigate the level of adherence to the advice generated by the SMSS, we separated between adherence to action feedback (supplied after the registration of a first measurement when further action was required) and adherence to the day conclusion (supplied when no further actions were required after the first registration of a day or when a second and final measurement was registered on the same day).

Action feedback could only appear in case the newly registered creatinine value was higher than the previous ones and required an additional measurement to confirm the first. In these cases, the feedback system of the SMSS showed an orange or red traffic light with the corresponding advice to repeat the measurement. From the SMSS-logged data, we selected those cases where a second measurement was requested and checked whether the concerning patients indeed measured and registered a second creatinine value on the same day.

To study adherence to the day conclusion, we only considered the cases in which patients again were confronted with an orange or a red traffic light. In case of a request to perform another measurement the next day (orange traffic light), the SMSS-logged data were checked to see whether the requested action was indeed performed. In case of a request to contact the hospital (red traffic light), patient hospital records were searched for telephonic and outpatient contacts on dates following the concerning feedback.

Sensitivity Analyses

Two sensitivity analyses were performed to control for potential bias. First, as being hospitalized limits the possibility to keep up with requested measurement and registration frequency, the level of adherence was analyzed with and without patients that were hospitalized during the study. The second sensitivity analysis concerned the test measurements patients were requested to perform when opening a new bottle of strips. Many patients either did not perform test measurements or did not indicate them as such. To prevent test values to be mistakenly considered creatinine measurements, all values that were stored in the device memory were checked. Potential test values were discussed and decided upon by the two main authors. The following criteria were used: (1) the value was not registered in the SMSS, (2) the value differed from the previous and following value, (3) the value fell within the test value range that was set by the manufacturer (133-239 µmol/l), and (4) the value followed or was followed by at least one SMSS-registered value measured on that same day (measured shortly after one another according to device-logged data). After having thoroughly checked and discussed all potential test values, for 24 values it remained unclear whether they were test values or not. We therefore performed all analyses concerning the representativeness of registered creatinine data with and without these 24 values.

Furthermore, we compared our findings concerning patient self-monitoring creatinine to a more broadly used and well-accepted form of patient monitoring, being self-monitoring of blood pressure. For this purpose, we performed two analyses with the self-monitored and self-reported blood pressure measurements in our study population. First, we looked at adherence to the blood pressure measurement protocol using the same procedure as for creatinine: number of days with measurements versus number of requested measurement days per patient per phase. The requested frequency of blood pressure measurements was equal to the measurement frequency of creatinine. As many patients used other blood pressure devices than the device we supplied for the study, we could not determine adherence to the measurement protocol in a reliable way. We therefore only could assess adherence to the registration protocol by comparing the number of registered blood pressure measurements (SMSS-logged data) with the requested measurement frequency.

Second, we looked at correspondence between measured and registered blood pressures. To prevent potential bias and misinterpretation, we (1) only used measurements from days with an equal number of measurements and registrations and (2) calculated a mean arterial pressure (MAP, [(2×diastolic)+systolic]/3) for both measured and registered blood pressures per day. Using all cases of noncorrespondence between measured and registered MAP, an overall MAP was calculated per patient for both blood pressures that were actually measured and for blood pressures that were registered in the SMSS. A paired *t* test was performed comparing these means.

Results

Participants

Within the period of inclusion, in total 217 patients received a kidney transplant of which 155 were considered eligible for participation. The main reasons for ineligibility were insufficient mastery of the Dutch language (25/62, 40%) and no access to a computer or limited computer skills (16/62, 26%). In total, 119 patients (76.7%) signed an informed consent. The main reason for not wanting to participate was the anticipated burden of self-monitoring (17/36, 47%). A total of 65 patients were randomized to the intervention group. After randomization, 3 patients dropped out because of graft dysfunction, death, and cancellation of transplantation (none was study related). Before starting to self-monitor kidney function at home, 4 patients canceled their participation because they reported having little trust in the creatinine device, experienced difficulties when logging into the SMSS, experienced business rush, or had a worsened condition post transplantation. In total, 58 patients were supplied with a creatinine and blood pressure device of which 4 never performed any measurement.

Table 1. Baseline patient characteristics.

Characteristics[a]	Received devices (n=58)	Used devices and SMSS during all phases (n=43)	*P*
Sex, male, n (%)	37 (64)	27 (63)	*P*>.99
Age at transplantation, mean (SD)	51.6 (14)	52.5 (15)	*P*=.40
Living together or married, n (%)	43 (74)	33 (77)	*P*=.50
Children, yes, n (%)	39 (67)	29 (67)	*P*=.76
Educational level, n (%)			*P*=.25
Low	22 (38)	14 (33)	
Middle	16 (28)	12 (28)	
High	20 (34)	17 (39)	
Paid job, yes, n (%)	31 (53)	23 (53)	*P*>.99
Origin, native, n (%)	53 (91)	41 (95)	*P*=.10
Former transplantation, n (%)	6 (10.3)	4 (9.3)	*P*=.64
Dialysis dependence before transplantation, n (%)	26 (44.8)	21 (49)	*P*=.14
Living transplantation, n (%)	50 (86.2)	38 (89)	*P*=.52
Kidney function (eGFR), mean (SD)	49 (16.1)	50 (15)	*P*=.34

[a]For a few patients, data on marital status and education were missing. These data were imputed in SPSS using multiple imputation (10 imputations).

To study the level of adherence to requested measurement frequency, we included patients of whom measured values were available for at least one complete study phase (N=48). To study the reliability of registered data and adherence to system feedback, we included patients who performed and registered measurements during all study phases (N=43). The flowchart in Figure 2 gives a stepwise overview of the patient flow and for which selection of patients a specific analysis was performed.

Patient characteristics are shown in Table 1. The mean age of participants was 52 and 53 years for patients who received the monitoring devices (N=58) and patients who performed and registered measurements during all study phases (N=43), respectively. The number of patients with both a low and a high educational level was slightly higher than in the average Dutch population [36]. Almost 90% of our participants (50/58) had received a kidney from a living donor, while the ratio of transplantations with living versus postmortem kidneys was about equal in our center during the period of inclusion. This discrepancy is mainly due to a higher percentage of ineligibility among recipients of a postmortem versus living kidney: 51% versus 16%, respectively.

Figure 2. Study flowchart.

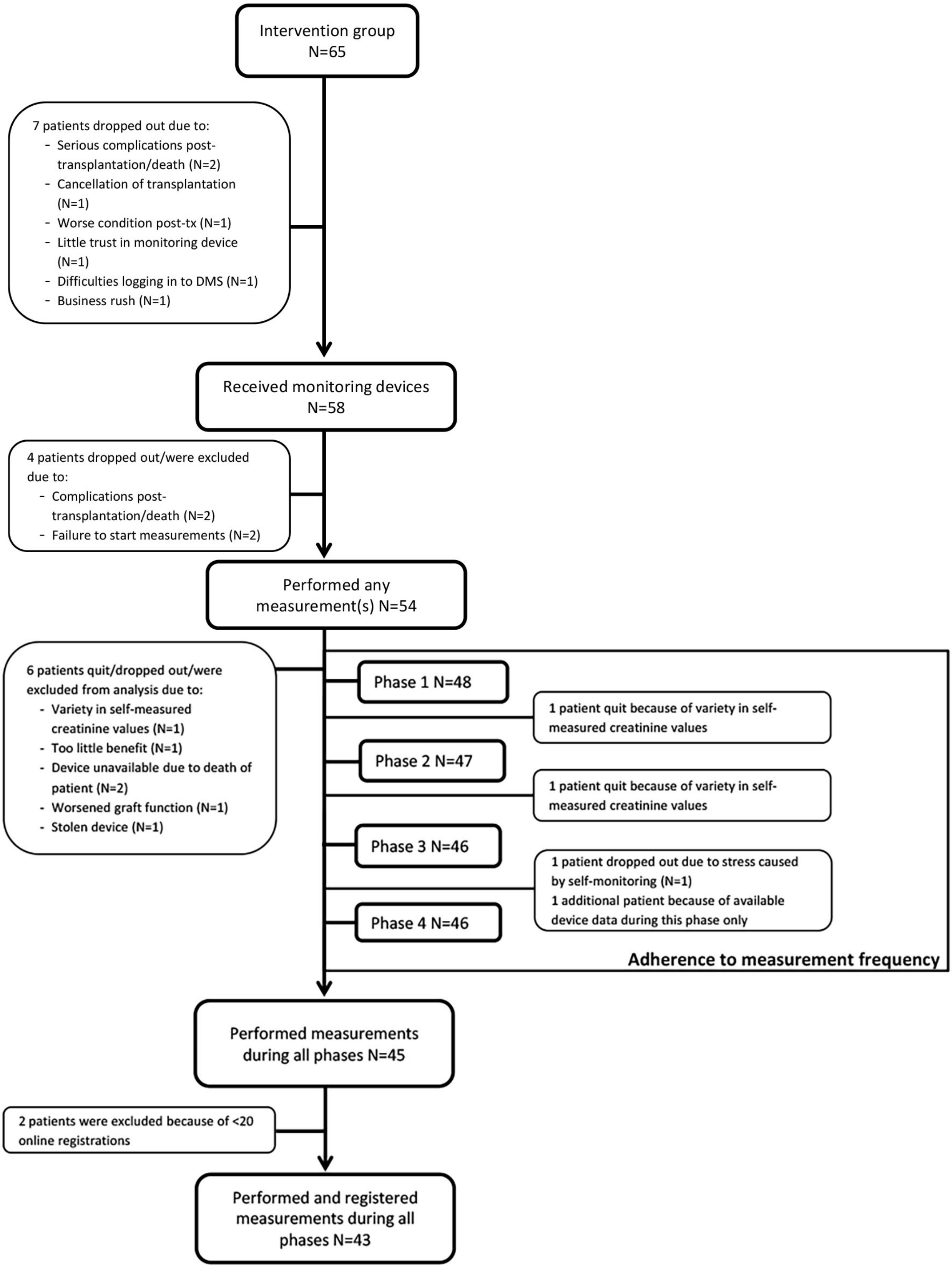


Figure 3. Patient adherence to requested measurement and Self-Management Support System (SMSS) registration frequency per study phase. The x-axis refers to the different measurement frequencies requested throughout the study for both performed and registered measurements (daily, every other day, twice a week, and weekly in phases 1, 2, 3, and 4, respectively), and the y-axis shows the percentage of adherent and overadherent and nonadherent and extremely nonadherent patients.

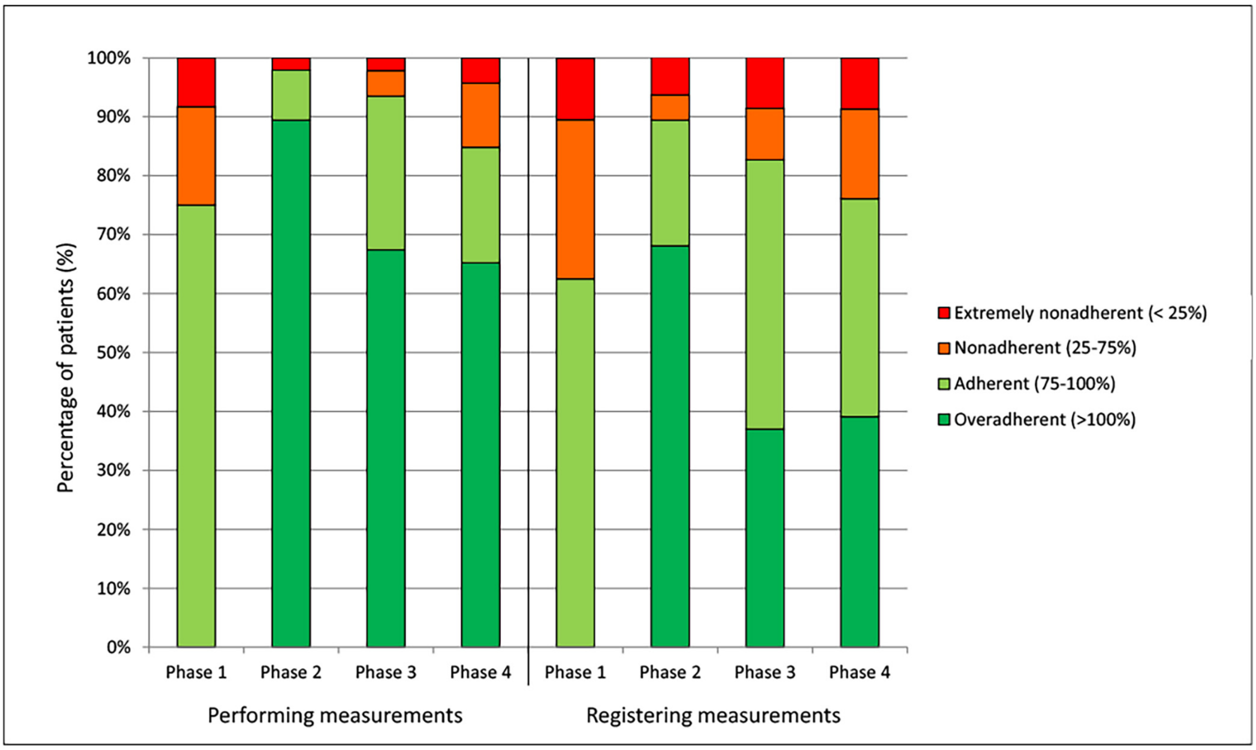
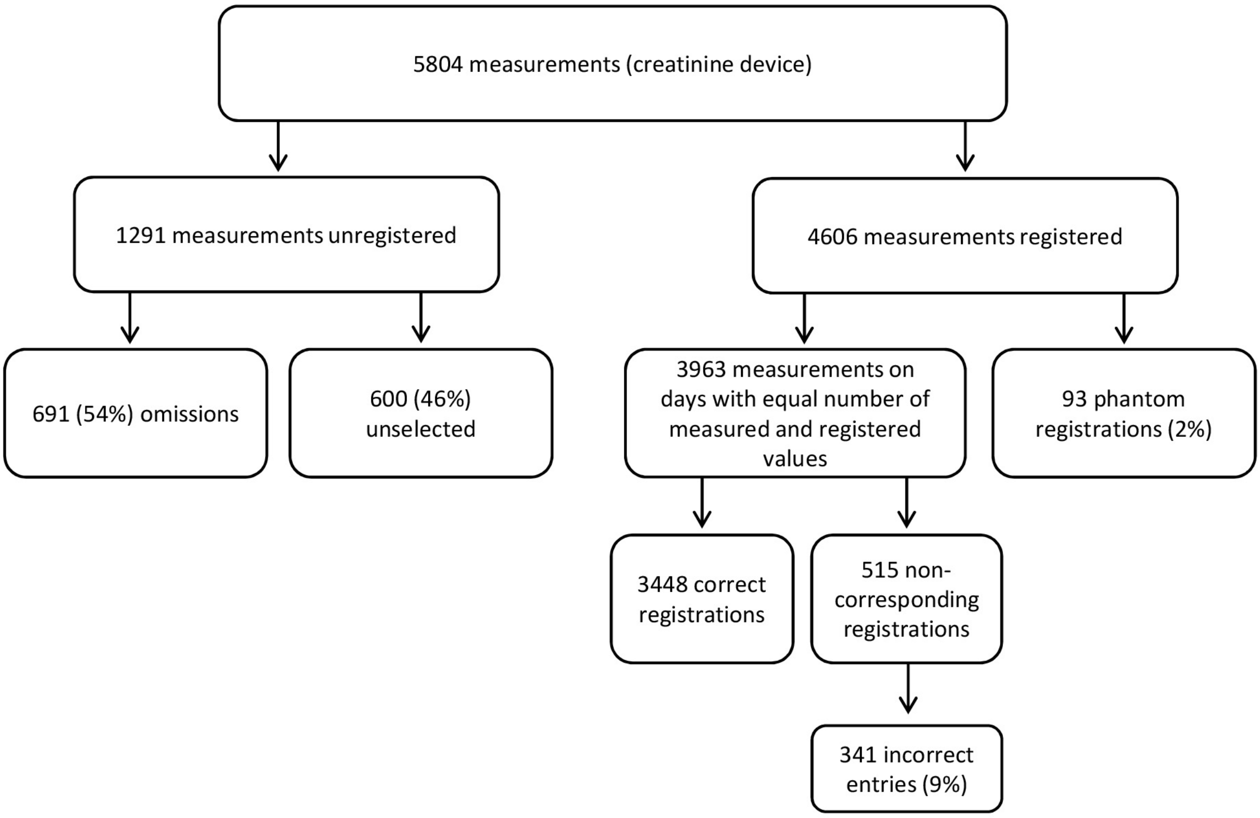


Figure 4. Reliability of creatinine values registered in the disease management system. The unregistered (1291) and registered values (4606) do not add up to the total number of measurements in the device (5804) due to the presence of phantom values.



No differences were found between patients who received the monitoring devices and patients who performed and registered measurements during all study phases for the characteristics we measured at baseline (see Table 1).

Adherence to Requested Measurement Frequency

Adherence to requested measurement frequency according to device-logged data (did patients perform the requested number of measurements?) and adherence to SMSS-logged data (if patients performed a measured, did they register the requested number of measurements?) are shown in Figure 3. During phase 1 (daily measurements), adherence to the requested number of creatinine measurements was lowest with 75% (36/48). Subsequently, adherence rose to over 90% during phases 2 (46/47, measuring every other day) and 3 (43/46, measuring twice a week), and then decreased to 85% during phase 4 (39/46, measuring weekly). In total, 4 patients performed less than 75% of the requested measurements throughout two or more phases. For registration of the requested number of creatinine measurements this same pattern is shown, although the percentage of nonadherent patients is higher during all phases. In total, 8 patients were nonadherent during two or more phases regarding registration of the requested number of measurements.

To control for the potential influence of hospitalization on the level of adherence to measurement and registration protocol, we repeated our analysis with all hospitalized patients excluded. The total number of hospitalized patients was 11 during phase 1, 3 during phase 2, 7 during phase 3, and 9 during phase 4. Excluding these patients did not change our initial findings.

Moment of Registration

When looking at the date of measurement versus the date of registration of measurements, a mean delay of 4 days (SD 10) was found. The level of delay varied from 1 to 81 days. A total of 7 patients (7/47, 15%) always registered their test results on the day of measurement and 15 patients (31%) had an overall mean delay of less than 1 day between measurement and registration. In total, 22 patients (46%) had an overall mean delay of more than 3 days, ranging up to a mean difference of 29 days between the date of measurement and the date of registration. One could hypothesize that patients do not feel the need to register their measurement if their level of creatinine is stable. However, the feedback that was generated by the SMSS for measurements registered on the day of measurement versus measurements that were registered with delay did not differ: in both situations, patients were requested to repeat the measurement in about 7% of all registrations.

The x-axis refers to the different measurement frequencies requested throughout the study for both performed and registered measurements (daily, every other day, twice a week, and weekly in phases 1, 2, 3, and 4, respectively), and the y-axis shows the percentage of adherent and overadherent and nonadherent and extremely nonadherent patients.

Reliability of Registered Data

Of the 43 patients included in the reliability analysis, the total number of values stored in the creatinine devices was 5779 and the total number of values registered in the SMSS was 4606. To investigate correspondence between measurement and registration, only days with an equal number of measurements in the device and registrations in the SMSS were selected. Total number of measurements performed and registered on these days was 3963. Figure 4 gives an overview of the reliability of the registered data, showing that 3448 (87.00%) of these values were registered correctly.

Noncorresponding Registrations

In 13% (515/3963) of all creatinine registrations, the registered value did not correspond to the value that was measured on that day. In 174 cases, we could determine the origin of the difference (eg, wrong combination of date and measured value, typo, rounding off). The remaining 341 registrations (9% of all registered values) were used for further analysis. Median number of noncorresponding values per patient was 3 (Interquartile range, IQR 8). Overall, 11 out of 43 patients (25%) made no mistakes at all, while another 11 patients made more than 10 mistakes. In total, 2 patients had an extremely high number of noncorresponding registrations. The first one had 52 noncorresponding registrations, half of which were found to be exactly 10, 20, 30, or 40 µmol/l lower than what was actually measured. In total, 83% (43/52) of his noncorresponding entries were lower than what was measured. The other patient registered 92 noncorresponding values, which were lower than the actual measured values in 93% (86/92) of his cases.

In case of noncorrespondence, the difference between measured and registered ranged from 1 to 73 mmol/l with a median of 9 mmol/l (IQR 13). The noncorresponding registrations were significantly lower than the actual measured ones: 123 mmol/l (SD 28) versus 130 mmol (SD 33), respectively (t_{340}=8.7, $P \leq .001$).

Phantom Values

In total, 93 phantom values were found, which was 2.02% of all registered values (N=4606). In total, 30% of patient (13/43) registered at least one phantom value, 14% (6/43) registered seven or more. Of the phantom values, 20 resembled the measurements of surrounding days, which would suggest these phantom values were only registered to adhere to the registration protocol. However, this appeared not to be the case as 16 of these 20 resembling phantom values were registered by a single patient who would already have been overly adherent without these phantom values. Overall, 3 patients registered creatinine values during the months where no measurements were logged in the device. For example, one patient quit measuring in February, but registered three measurements during March and April. Phantom values were significantly lower than actually measured ones, respectively 107 (SD 26) and 123 (15) mmol/l (t_{11}=3.9, P=.003).

Representation of Registered Creatinine Values

Of the 5779 measurements found in the creatinine device-logged data, 1300 values (22.49%) were not registered in the SMSS. In 700 cases, one or more measurements were performed on a given date, but no value was registered in the SMSS (ie, omission). Number of omissions per patient ranged from 0 to 145; 5 patients had no omissions at all and 11 patients had omitted 20 values or more. Median number of omissions was

8. The omitted values were significantly higher than the registered ones (mean of 139 [SD 31] vs 130 [SD 32] µmol/l, respectively (t_{42}=−3.8, P ≤.001).

In several cases, more measurements were performed per day than values were registered. In these so-called measurement series, the number of performed measurements ranged from 2 to 8 with a median of 2 (IQR 1) per day. The total number of values that was measured within a series of measurements but was not registered in the SMSS was 600. The mean of the creatinine values that were both stored in the device and registered in the SMSS was significantly lower than the mean of the creatinine values that were stored in the device but not registered in the SMSS (unselected for registration): 137 (SD 35) vs 143 µmol/l (SD 36), respectively (t_{42}=−2.5, P=.02).

Repeating both analyses with the 24 values included that could either be test values or regular measurements did not change our findings (data not shown).

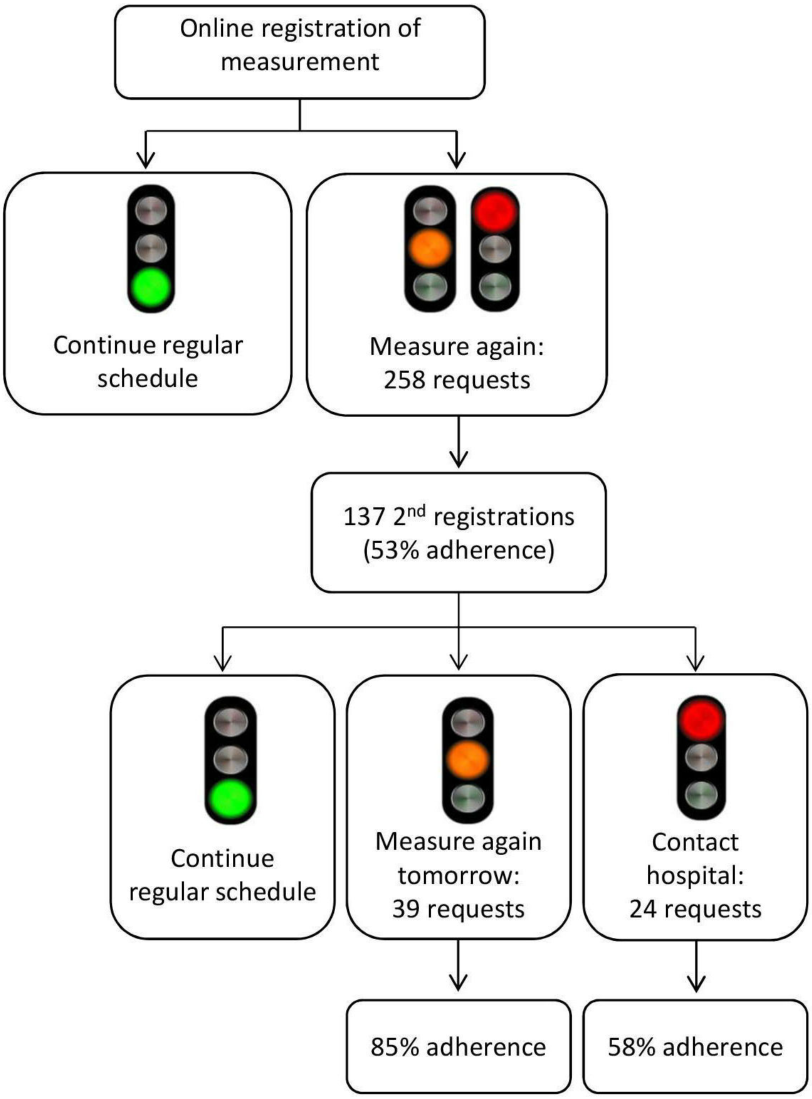
Adherence to System Feedback

Results are given separately for adherence to the action code (requesting patients to perform a second measurement directly) and adherence to the day conclusion (feedback that only appeared in case a second measurement was requested and registered). An overview of the feedback procedure and level of adherence to the different kinds of feedback is shown in Figure 5.

Adherence to Action Code

In 258 cases, patients were requested to perform a second measurement directly. In 137 cases (53%), patients actually performed and registered a second measurement. In 85 cases (33%), date of registration differed from the date of measurement, suggesting that delayed registration was the main reason for not adhering to feedback, as this feedback was no longer up-to-date when shown to the patient. In 14 cases (5%), multiple measurements were performed in advance of registration. As these multiple measurements were probably representative of actual creatinine level already, patients might not have felt the need to perform another one.

Figure 5. Patient adherence to system feedback.



Adherence to Day Conclusion

In 39 cases, patients were requested to measure again the next day. In 33 cases (85%), this feedback was adhered to. In 3 cases (8%), the measurements were registered several days after measurement (delayed registration), suggesting that the feedback to measure again the next day was no longer up-to-date when shown to these patients. In one case, the requested measurement was performed, but was not registered in the SMSS.

The advice to contact the hospital was given 24 times, which was followed up 14 times (58%). In the remaining 10 cases (concerning 10 individual patients), measurements were registered with several days' delay. As feedback was no longer up-to-date when shown to the concerning patients, this suggests delayed registration was the main reason for not adhering to the feedback to contact the hospital. If only cases with registration on the day of measurement were taken into account, adherence to contacting the hospital was 100%.

Sensitivity Analysis: Adherence to and Reliability of Blood Pressure Measurements

In total, 31 blood pressure devices could be read out. The total number of values found in the 31 available pressure devices was 4917, and the total number of values registered in the SMSS was 5637. The higher number of registered than measured blood pressures is due to patients using multiple blood pressure devices during study participation, while data of only one device were available. Adherence to registration of blood pressure measurements was comparable to creatinine registrations with 70% (34/48), 88% (41/47), 87% (40/46), and 81% (37/46) of patients registering >75% of the requested number of measurements during phases 1, 2, 3, and 4, respectively. Although the percentage of registrations not corresponding to the measured MAP was comparable to what we found for creatinine (14% vs 13%, respectively), we could not replicate the significant difference between registered and actually measured creatinine for blood pressure: 97 mmHg (SD 2) versus 96 mmHg (SD 9) for registered and measured MAPs, respectively (t_3=.20, P=.84).

Discussion

Principal Findings

Self-monitoring kidney function and blood pressure at home offers important advantages for patients after kidney transplantation. However, the value and safety of self-monitoring depend on how well patients actually adhere to their self-monitoring tasks, the reliability of the test results they report, and whether they take appropriate actions based on their measurements. This study showed that the level of adherence was generally good. Well above 90% of all patients performed the requested number of measurements during months 2-4 post transplantation. Adherence was lower during the first month when more measurements were requested and during months 5-12 post transplantation when less measurements were requested, with about 75% and 85% of patients adhering to the requested number of measurements, respectively. Overall adherence to registration of measurements was about 10% lower than adherence to performance of measurements during all phases. Two studies reporting on the level of adherence to monitoring vital signs after lung transplantation found similar percentages of adherence being above 80% for the entire study period [24,26]. For self-monitoring blood pressure, patients with uncontrolled hypertension were shown to be adherent for about 73% of the entire study period [25,29]. In both studies, level of adherence was highest in the first few weeks and declined gradually over time. In sum, mean level of adherence that has been found in this study corresponds to percentages that have previously been reported. In contrast, we did not find the highest levels of adherence in the first period. This may have been due to a strenuous measurement protocol. Patients had to measure every day in the first month. In these first weeks when patients have to recover and have to get used to life post transplantation, performing measurements in such a high frequency might be too burdensome. Furthermore, in this first period, face-to-face visits were not yet replaced by telephonic consults and patients therefore visited the hospital at least weekly to monitor early signs of graft failure. Due to this high frequency of visits, patients may have felt a reduced need to perform measurements at home, as they did not have to rely on these measurements. The latter may also be an explanation for nonadherence during the whole study period.

Furthermore, for self-monitoring to be a safe alternative to regular face-to-face follow up, patient-reported test results need to be accurate. In this study, approximately 90% of both creatinine and blood pressure measurements were registered correctly in the SMSS. This percentage corresponds to what has previously been described for patient-reported blood pressure [29,37] and anticoagulation [34] and is much higher than has been observed for patient-reported levels of blood glucose. A study by Kalergis and colleagues [30], for example, showed that slightly over half of the total group of patients with either diabetes type 1 or 2 was considered very reliable in their reporting. For patients with diabetes type 2 and for pregnant women self-monitoring blood glucose, some studies even showed that the majority of patient-reported data was unreliable [31,32].

In cases of noncorrespondence between measured and actually registered values, values that were eventually registered in the SMSS were significantly lower than those actually measured. These results seem to suggest that patients select, alter, or add values in such a way that their creatinine profile looks more positive. This corresponds to what has been found in a population of patients with thrombosis, where the percentage of time when patients' level of anticoagulation was within the desired range was significantly higher when using patient-reported data compared with data stored in the device [34]. For patients with diabetes or hypertension, it was found that inaccurate reporting increased with increasing levels of blood glucose [31] or blood pressure [37]. Why patients report values that look better than the actual measured values or add nonexistent measurements has not yet been fully clarified. For diabetes, it has been suggested that patients report false glucose levels due to a feeling of guilt for not having achieved glycemic goals [32] or add phantoms values in an attempt to fill up logbooks and satisfy their health care providers [30]. Both situations seem to represent an attempt to be a "good" patient.

However, altering and selecting data that are not representative of the actual clinical situation or adding phantom values in any case may be dangerous. This can lead to suboptimal treatment and, eventually, to worsened patient outcomes [30,37]. In a study by Kendrick and colleagues, it was found indeed that women with pregnancy-derived diabetes received suboptimal treatment due to a large difference between their reported glucose values and what they had measured [32]. Results of another study showed that diabetic patients who were more reliable in their reporting had a significantly better glycemic control. It was suggested that this may be due to the ability of clinicians to adjust therapy more precisely if measurements are reported accurately [30]. To prevent incorrect reporting, it has been recommended to rely on the memory capacity of measurement devices, preferably by using devices that can transfer data automatically [30,31,33,34].

Besides eliminating the occurrence of both intentional and unintentional errors, the automatic transfer of data offers a solution for the observation that patients seem to save up their measurements before registering them. Many patients saved up their measurements over several days or even weeks to register them all at once. More than one-third of our participants displayed a mean delay of 5 or more days between measurement and registration of data. This is alarming as frequent monitoring and taking immediate action in case of early signs of graft failure is vital to prevent or diminish damage to the kidney transplant. An explanation for saving up measurements before registering them might be that the measured creatinine values remained stable. However, patients seemed to postpone registration regardless of the stability of their kidney function. Indeed, postponement of registration appeared to be the main reason why patients had not followed up the advice to contact the hospital when creatinine levels had alarmingly increased by over 15%. Patients' perception of these significant increases could have been influenced by the fact that the innovative device that was used during this study tended to be less accurate than hospital laboratory measurements [35]. As a consequence, patients might have been inclined to attribute sudden increases in level of creatinine to a technical imprecision of the device.

Strengths and Limitations

To the best of our knowledge, this is the first study to assess adherence to a protocol of self-monitoring creatinine and to investigate the accuracy and reliability of patient-generated creatinine data. Enabling patients to self-monitor kidney function at home would have important advantages, especially for patients living in remote areas. There are, however, some limitations that must be considered when interpreting our findings. First, the study was conducted at a single institute. As each hospital has its own way of delivering care, results might be different when conducted in other institutions. Our findings do, however, resemble what has previously been found in other disease populations. Second, participation in this study was voluntary and we selected patients with access to the Internet. Therefore, it is possible that our patients had an above-average motivation to self-monitor. It is therefore likely that the current findings provide a conservative estimate of the true incidence of nonadherence and inaccuracy. In line with this, patients reported very strong intentions to engage in self-monitoring both at the start and after 4 months into the trial [38]. Their intention was found to be especially associated with their overall affective reaction toward using the system [38]. A considerable number of eligible patients had to be contacted to inform whether they were interested in study participation instead of giving informed consent immediately. However, the limited variance found in the level of intention to engage in self-monitoring suggests that patients who had to be contacted were not more hesitant to engage in self-monitoring than patients who provided immediate informed consent.

Finally, the way adherence and nonadherence were determined is arbitrary to some extent. As the importance of frequent and very frequent monitoring differs per subpopulation and parameter of interest, no gold standard for what can be considered adherent is available. High blood pressure, for example, needs to be present over a longer period of time before becoming detrimental, while an increasing level of creatinine can be indicative of a rejection episode, leading to irreversible damage or even loss of the transplanted kidney if not quickly noticed.

Implications

This study shows that the level of adherence to a protocol of self-monitoring creatinine in the first year after kidney transplantation was generally good, although adherence declined over time. In addition, our results suggest that measuring every day in the first period after transplantation might be too burdensome. Furthermore, 90% of data were shown to be accurately reported. In line with previous findings, however, several patients reported more favorable data than they actually measured. This suggests that some patients might be inclined to select more favorable values for registration, which could leave early signs of graft failure unnoticed. Additionally, the majority of patients did not register their measured values on the day of measurement, but saved up measurements over several days to register them all at once. This so-called delayed registration was the main reason for patients not having followed up the advice to contact the hospital in case of a significantly increased level of creatinine.

This study is part of a larger project in which the safety and usability of self-monitoring kidney function after transplantation supported by an SMSS are investigated. Our results showing that patients seem inclined to select more favorable creatinine values for registration and to postpone registration suggest a challenge to the safety of self-monitoring. This should be well considered when designing self-monitoring care systems, for example by ensuring that self-measured data are transferred automatically to an SMSS. Using devices that can transfer data automatically and providing active feedback to patients (eg, by sending text messages or emails) instead of having patients to log on to a website will eliminate the issues of data selection and delayed registration, and as such contribute to the safety of self-monitoring kidney function after transplantation.

Acknowledgments

This work was supported by a grant from the Netherlands Organization for Health Research and Development (ZonMw) and the Dutch Kidney Foundation.

Conflicts of Interest

None declared.

References

1. Traynor J, Mactier R, Geddes CC, Fox JG. How to measure renal function in clinical practice. BMJ 2006 Oct 07;333(7571):733-737 [FREE Full text] [doi: 10.1136/bmj.38975.390370.7C] [Medline: 17023465]
2. Fernández-Fresnedo G, Palomar R, Escallada R, de Francisco M, Cotorruelo JG, Zubimendi JA, et al. Hypertension and long-term renal allograft survival: effect of early glomerular filtration rate. Nephrol Dial Transplant 2001;16(Suppl 1):105-109. [Medline: 11369835]
3. Hillebrand U, Suwelack BM, Loley K, Lang D, Reuter S, Amler S, et al. Blood pressure, antihypertensive treatment, and graft survival in kidney transplant patients. Transpl Int 2009 Nov;22(11):1073-1080 [FREE Full text] [doi: 10.1111/j.1432-2277.2009.00922.x] [Medline: 19624495]
4. Mange KC, Cizman B, Joffe M, Feldman HI. Arterial hypertension and renal allograft survival. JAMA 2000 Feb 02;283(5):633-638. [Medline: 10665703]
5. Opelz G, Döhler B, Collaborative Transplant Study. Improved long-term outcomes after renal transplantation associated with blood pressure control. Am J Transplant 2005 Nov;5(11):2725-2731 [FREE Full text] [doi: 10.1111/j.1600-6143.2005.01093.x] [Medline: 16212633]
6. Agarwal R, Bills J, Hecht TJ, Light RP. Role of home blood pressure monitoring in overcoming therapeutic inertia and improving hypertension control: a systematic review and meta-analysis. Hypertension 2011 Jan;57(1):29-38 [FREE Full text] [doi: 10.1161/HYPERTENSIONAHA.110.160911] [Medline: 21115879]
7. Bloomfield H, Krause A, Greer N, Taylor BC, MacDonald R, Rutks I, et al. Meta-analysis: effect of patient self-testing and self-management of long-term anticoagulation on major clinical outcomes. Ann Intern Med 2011 Apr 05;154(7):472-482. [doi: 10.7326/0003-4819-154-7-201104050-00005] [Medline: 21464349]
8. Bosworth H, Olsen MK, Grubber JM, Neary AM, Orr MM, Powers BJ, et al. Two self-management interventions to improve hypertension control: a randomized trial. Ann Intern Med 2009 Nov 17;151(10):687-695 [FREE Full text] [doi: 10.7326/0003-4819-151-10-200911170-00148] [Medline: 19920269]
9. Cappuccio FP, Kerry SM, Forbes L, Donald A. Blood pressure control by home monitoring: meta-analysis of randomised trials. BMJ 2004 Jul 17;329(7458):145 [FREE Full text] [doi: 10.1136/bmj.38121.684410.AE] [Medline: 15194600]
10. Gardiner C, Longair I, Pescott MA, Erwin H, Hills J, Machin SJ, et al. Self-monitoring of oral anticoagulation: does it work outside trial conditions? J Clin Pathol 2009 Feb;62(2):168-171 [FREE Full text] [doi: 10.1136/jcp.2008.059634] [Medline: 19181634]
11. Glynn L, Murphy AW, Smith SM, Schroeder K, Fahey T. Self-monitoring and other non-pharmacological interventions to improve the management of hypertension in primary care: a systematic review. Br J Gen Pract 2010 Dec;60(581):e476-e488 [FREE Full text] [doi: 10.3399/bjgp10X544113] [Medline: 21144192]
12. Heneghan C, Garcia-Alamino JM, Spencer EA, Ward AM, Perera R, Bankhead C, et al. Self-monitoring and self-management of oral anticoagulation. Cochrane Database Syst Rev 2016 Jul 05(7):CD003839. [doi: 10.1002/14651858.CD003839.pub3] [Medline: 27378324]
13. Margolis K, Asche SE, Bergdall AR, Dehmer SP, Groen SE, Kadrmas HM, et al. Effect of home blood pressure telemonitoring and pharmacist management on blood pressure control: a cluster randomized clinical trial. JAMA 2013 Jul 03;310(1):46-56 [FREE Full text] [doi: 10.1001/jama.2013.6549] [Medline: 23821088]
14. McManus RJ, Mant J, Haque MS, Bray EP, Bryan S, Greenfield SM, et al. Effect of self-monitoring and medication self-titration on systolic blood pressure in hypertensive patients at high risk of cardiovascular disease: the TASMIN-SR randomized clinical trial. JAMA 2014 Aug 27;312(8):799-808. [doi: 10.1001/jama.2014.10057] [Medline: 25157723]
15. Uhlig K, Patel K, Ip S, Kitsios GD, Balk EM. Self-measured blood pressure monitoring in the management of hypertension: a systematic review and meta-analysis. Ann Intern Med 2013 Aug 06;159(3):185-194. [doi: 10.7326/0003-4819-159-3-201308060-00008] [Medline: 23922064]
16. Barnard KD, Young AJ, Waugh NR. Self monitoring of blood glucose - a survey of diabetes UK members with type 2 diabetes who use SMBG. BMC Res Notes 2010 Nov 22;3:318 [FREE Full text] [doi: 10.1186/1756-0500-3-318] [Medline: 21092171]
17. Jaana M, Paré G. Home telemonitoring of patients with diabetes: a systematic assessment of observed effects. J Eval Clin Pract 2007 Apr;13(2):242-253. [doi: 10.1111/j.1365-2753.2006.00686.x] [Medline: 17378871]
18. Jones MI, Greenfield SM, Bray EP, Baral-Grant S, Hobbs FD, Holder R, et al. Patients' experiences of self-monitoring blood pressure and self-titration of medication: the TASMINH2 trial qualitative study. Br J Gen Pract 2012 Feb;62(595):e135-e142 [FREE Full text] [doi: 10.3399/bjgp12X625201] [Medline: 22520791]

19. Sicotte C, Paré G, Morin S, Potvin J, Moreault M. Effects of home telemonitoring to support improved care for chronic obstructive pulmonary diseases. Telemed J E Health 2011 Mar;17(2):95-103. [doi: 10.1089/tmj.2010.0142] [Medline: 21214399]
20. Braun S, Spannagl M, Völler H. Patient self-testing and self-management of oral anticoagulation. Anal Bioanal Chem 2009 Mar;393(5):1463-1471. [doi: 10.1007/s00216-008-2225-3] [Medline: 18568340]
21. Cromheecke M, Levi M, Colly LP, de Mol BJ, Prins MH, Hutten BA, et al. Oral anticoagulation self-management and management by a specialist anticoagulation clinic: a randomised cross-over comparison. Lancet 2000 Jul 08;356(9224):97-102. [doi: 10.1016/S0140-6736(00)02470-3] [Medline: 10963245]
22. Ryan F, Byrne S, O'Shea S. Randomized controlled trial of supervised patient self-testing of warfarin therapy using an internet-based expert system. J Thromb Haemost 2009 Aug;7(8):1284-1290 [FREE Full text] [doi: 10.1111/j.1538-7836.2009.03497.x] [Medline: 19496921]
23. van Lint CL, van der Boog PJ, Wang W, Brinkman W, Rövekamp TJ, Neerincx MA, et al. Patient experiences with self-monitoring renal function after renal transplantation: results from a single-center prospective pilot study. Patient Prefer Adherence 2015;9:1721-1731 [FREE Full text] [doi: 10.2147/PPA.S92108] [Medline: 26673985]
24. Finkelstein SM, Snyder M, Edin-Stibbe C, Chlan L, Prasad B, Dutta P, et al. Monitoring progress after lung transplantation from home-patient adherence. J Med Eng Technol 1996;20(6):203-210. [Medline: 9029392]
25. Kerby TJ, Asche SE, Maciosek MV, O'Connor PJ, Sperl-Hillen JM, Margolis KL. Adherence to blood pressure telemonitoring in a cluster-randomized clinical trial. J Clin Hypertens (Greenwich) 2012 Oct;14(10):668-674 [FREE Full text] [doi: 10.1111/j.1751-7176.2012.00685.x] [Medline: 23031143]
26. Morlion B, Knoop C, Paiva M, Estenne M. Internet-based home monitoring of pulmonary function after lung transplantation. Am J Respir Crit Care Med 2002 Mar 01;165(5):694-697. [doi: 10.1164/ajrccm.165.5.2107059] [Medline: 11874816]
27. Patton SR. Adherence to glycemic monitoring in diabetes. J Diabetes Sci Technol 2015 May;9(3):668-675 [FREE Full text] [doi: 10.1177/1932296814567709] [Medline: 25591853]
28. Vincze G, Barner JC, Lopez D. Factors associated with adherence to self-monitoring of blood glucose among persons with diabetes. Diabetes Educ 2004;30(1):112-125. [doi: 10.1177/014572170403000119] [Medline: 14999899]
29. Cheng C, Studdiford JS, Chambers CV, Diamond JJ, Paynter N. The reliability of patient self-reported blood pressures. J Clin Hypertens (Greenwich) 2002;4(4):259-264 [FREE Full text] [Medline: 12147928]
30. Kalergis M, Nadeau J, Pacaud D, Yared Z, Yale JF. Accuracy and reliability of reporting self-monitoring of blood glucose results in adults with type 1 and type 2 diabetes. Can J Diabetes 2006 Jan;30(3):241-247 [FREE Full text] [doi: 10.1016/S1499-2671(06)03006-1]
31. Kazlauskaite R, Soni S, Evans AT, Graham K, Fisher B. Accuracy of self-monitored blood glucose in type 2 diabetes. Diabetes Technol Ther 2009 Jun;11(6):385-392. [doi: 10.1089/dia.2008.0111] [Medline: 19459768]
32. Kendrick JM, Wilson C, Elder RF, Smith CS. Reliability of reporting of self-monitoring of blood glucose in pregnant women. J Obstet Gynecol Neonatal Nurs 2005;34(3):329-334. [doi: 10.1177/0884217505276306] [Medline: 15890831]
33. Mengden T, Hernandez MR, Beltran B, Alvarez E, Kraft K, Vetter H. Reliability of reporting self-measured blood pressure values by hypertensive patients. Am J Hypertens 1998 Dec;11(12):1413-1417. [Medline: 9880121]
34. Quin J, Rogers LQ, Markwell S, Butler T, McClafferty R, Hazelrigg S. Home-anticoagulation testing: accuracy of patient-reported values. J Surg Res 2007 Jun 15;140(2):189-193. [doi: 10.1016/j.jss.2007.01.036] [Medline: 17397870]
35. van Lint CL, van der Boog PJ, Romijn FP, Schenk PW, van Dijk S, Rövekamp TJ, et al. Application of a point of care creatinine device for trend monitoring in kidney transplant patients: fit for purpose? Clin Chem Lab Med 2015 Sep 01;53(10):1547-1556. [doi: 10.1515/cclm-2014-0932] [Medline: 25719331]
36. Bierings H. CBS. 2013. Onderwijsniveau bevolking gestegen URL: https://www.cbs.nl/nl-nl/nieuws/2013/40/onderwijsniveau-bevolking-gestegen [accessed 2017-07-27] [WebCite Cache ID 6sGVBLBmT]
37. Johnson K, Partsch DJ, Rippole LL, McVey DM. Reliability of self-reported blood pressure measurements. Arch Intern Med 1999;159(22):2689-2693. [Medline: 10597759]
38. Wang W, van Lint CL, Brinkman WP, Rövekamp TJ, van Dijk S, van der Boog PJ, et al. Renal transplant patient acceptance of a self-management support system. BMC Med Inform Decis Mak 2017 May 08;17(1):58 [FREE Full text] [doi: 10.1186/s12911-017-0456-y] [Medline: 28482896]

Abbreviations

eGFR: estimated Glomerular Filtration Rate
IQ: interquartile range
LUMC: Leiden University Medical Center
MAP: mean arterial pressure
RCT: randomized controlled trial
SD: standard deviation
SMSS: Self-Management Support System

Edited by G Eysenbach; submitted 21.02.17; peer-reviewed by D Schatell, R Beyth, M Teimoori; comments to author 24.05.17; revised version received 29.06.17; accepted 30.06.17; published 26.09.17.

Please cite as:
van Lint C, Wang W, van Dijk S, Brinkman WP, Rövekamp TJM, Neerincx MA, Rabelink TJ, van der Boog PJM
Self-Monitoring Kidney Function Post Transplantation: Reliability of Patient-Reported Data
J Med Internet Res 2017;19(9):e316
URL: http://www.jmir.org/2017/9/e316/
doi:10.2196/jmir.7542
PMID:28951385

Original Paper

Putting the Focus Back on the Patient: How Privacy Concerns Affect Personal Health Information Sharing Intentions

Mohamed Abdelhamid[1], PhD; Joana Gaia[2], PhD; G Lawrence Sanders[2], PhD

[1]College of Business Administration, Department of Information Systems, California State University Long Beach, Long Beach, CA, United States
[2]Management Science and Systems, School of Management, University at Buffalo, Buffalo, NY, United States

Corresponding Author:
Mohamed Abdelhamid, PhD
College of Business Administration
Department of Information Systems
California State University Long Beach
1250 Bellflower Boulevard
Long Beach, CA, 90840
United States
Phone: 1 562 985 4993
Fax: 1 562 985 4993
Email: mohamed.abdelhamid@csulb.edu

Abstract

Background: Health care providers are driven by greater participation and systemic cost savings irrespective of benefits to individual patients derived from sharing Personal Health Information (PHI). Protecting PHI is a critical issue in the sharing of health care information systems; yet, there is very little literature examining the topic of sharing PHI electronically. A good overview of the regulatory, privacy, and societal barriers to sharing PHI can be found in the 2009 Health Information Technology for Economic and Clinical Health Act.

Objective: This study investigated the factors that influence individuals' intentions to share their PHI electronically with health care providers, creating an understanding of how we can represent a patient's interests more accurately in sharing settings, instead of treating patients like predetermined subjects. Unlike privacy concern and trust, patient activation is a stable trait that is not subject to change in the short term and, thus, is a useful factor in predicting sharing behavior. We apply the extended privacy model in the health information sharing context and adapt this model to include patient activation and issue involvement to predict individuals' intentions.

Methods: This was a survey-based study with 1600+ participants using the Health Information National Trends Survey (HINTS) data to validate a model through various statistical techniques. The research method included an assessment of both the measurement and structural models with post hoc analysis.

Results: We find that privacy concern has the most influence on individuals' intentions to share. Patient activation, issue involvement, and patient-physician relationship are significant predictors of sharing intention. We contribute to theory by introducing patient activation and issue involvement as proxies for personal interest factors in the health care context.

Conclusions: Overall, this study found that although patients are open to sharing their PHI, they still have concerns over the privacy of their PHI during the sharing process. It is paramount to address this factor to increase information flow and identify how patients can assure that their privacy is protected. The outcome of this study is a set of recommendations for motivating the sharing of PHI. The goal of this research is to increase the health profile of the patients by integrating the testing and diagnoses of various doctors across health care providers and, thus, bring patients closer to the physicians.

(J Med Internet Res 2017;19(9):e169) doi:10.2196/jmir.6877

KEYWORDS

information sharing; health information technology; privacy; personal health information

Introduction

Background

This study investigates critical factors that influence individuals' decisions to share their personal health information (PHI) electronically among health care providers. The paper focuses on individuals' information privacy concern, patient-physician relationship, trust in health care providers, and health-related factors such as patient activation and issue involvement. There is extensive literature investigating the behavior of people who seek information on the Web, yet there is only a modest amount of research studying what factors influence private health information sharing. The impetus for investigating the barriers to sharing health information was the 2009 Health Information Technology for Economic and Clinical Health (HITECH) Act. The passing of the HITECH Act marked a significant change in the appropriate protections and processes for sharing health information. With the signing of the HITECH Act of 2009, by President Obama, incentives and guidelines were established for health care providers for the use of Health Information Exchanges (HIEs). This Act provides a strong rationale for the development of HIEs. Indeed, it is recognized that exchanging patients' health information electronically improves the quality of care, reduces medical errors, and reduces medical costs [1-3]. However, patients' information cannot be shared, unless patients agree to share via an HIE. The value of HIE, therefore, is directly related to the relative ease of sharing among providers, payers, and patients [4]. There is a noticeable effort in literature trying to investigate the reasons affecting the willingness to share. Sharing patients' health information has been considered an urgency to promote investments in health care information technology [5]. Patients' decisions not to share may result in medical errors and undesired health outcomes. Our aspiration to understand the psychology behind patients' decision comes from our desire to address barriers to sharing and enhance motivators of sharing to help patients make better choices for their own health [5]. The findings of this paper will help health care stakeholders and policy makers enhance sharing of PHI to achieve better health care.

The motivation for this study comes from the willingness to provide practical and theoretical implications to the health care field for understanding factors affecting patients' intentions to share their PHI. Few studies have investigated ways to persuade patients to share their health information [6]; in fact, factors influencing patients' decisions related to sharing health information have not been studied thoroughly [7], and this study aims to address the gap in existing literature. Once this understanding is achieved, health care providers can promote motivational factors to help improve the sharing of PHI according to patients' needs and concerns and this will lead to better health outcomes and reduced medical costs for the entire population. In this study, the factors of interest are those that explain the decision to share or not to share PHI with health care providers.

Prior Literature

The literature that addresses attitudes toward sharing PHI is scarce and often characterized by studies with small sample sizes or only applicable to one group of the patient population [8,9]. Extant literature in attitudes toward sharing PHI usually addresses either the patients' willingness to share or barriers to sharing PHI.

Patients are typically very willing to share their PHI and have a positive attitude toward sharing practices [10-12], and this willingness is enhanced when potential privacy concerns are addressed [11,13]. Patients who trust their clinicians [14] and who can understand the health benefits brought by sharing PHI practices [6] will also be willing to share their PHI.

The studies that report barriers to sharing PHI point out a variety of factors that hinder participation in HIEs or electronic sharing of PHI. The foremost barrier encountered in the studies is related to the privacy and security of PHI. Patients are apprehensive about who will have access to their PHI [6,8,14], how it will be used in HIEs [15], and the intentions of the PHI users [16]. Patients who have control over how much information they share and who they share it with are more prone to sharing [9,15,17]. Patients who do not have a perception of the benefits brought by PHI sharing are also less willing to share their information [8,18]. Low income, ethnic diversity [14], general health status, certain personality traits [19], and exiting medical conditions (eg, depression) [8] have all been identified as factors that hinder engaging in PHI sharing practices.

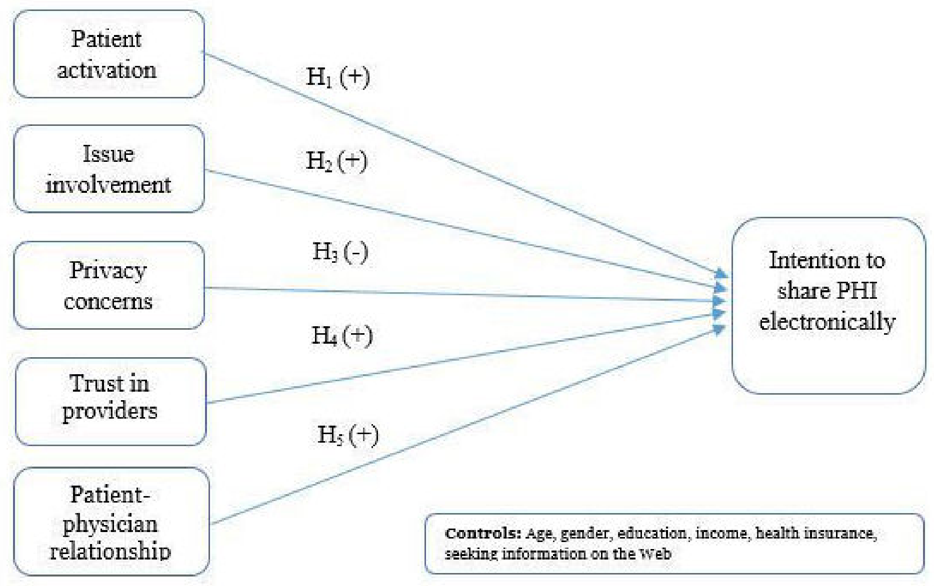
In summary, the majority of studies in the attitude toward sharing PHI arena have shown that patients encounter many barriers that thwart the above-mentioned willingness to share PHI. It is therefore critical to understand how consumers can be educated and how their concerns can be addressed to achieve higher sharing rates. Our ultimate goal is to help patients achieve optimal health outcomes while protecting their privacy.

Theoretical Foundation

This study investigates the impact of concern for information privacy, trust in providers, patient activation, issue involvement and patient-physician relationship on the intention to share PHI. The overall contribution of this study is to provide theoretical and practical insights to address the privacy and trust barriers in consent. Moreover, we use traits such as patient activation to predict patients' intentions to share their PHI. Unlike trust and privacy, these traits are not subject to changes at least in the short term and, thus, are stable and robust predictors of individuals' intentions.

The proposed model has been theoretically developed based on the extended privacy calculus model [20]. Many studies in the information systems field focused on understanding people's intentions and behavior toward information technology, and based their theoretical foundations on the theory of reasoned action and the theory of planned behavior [21,22]. However, the extended privacy calculus model is specifically intended to predict intention of information disclosure in the Internet environment, and thus is more appropriate for our study. The model proposes that an individual's intention to disclose information in Web-based transactions depends on their privacy concerns, trust in the system, and their personal interest in the context of the transaction.

Figure 1. Conceptual research model.



Research Model and Hypotheses

We adapt the extended privacy calculus model and use it in the context of health care information systems to predict individuals' intentions to share their PHI electronically among providers. We introduce patient activation and issue involvement predictors of personal interest to engage in the sharing behavior in the health care context, as seen in Figure 1.

Patient Activation

A person's decision to engage in a behavior depends on their motivation to engage in that behavior [23,24]. Motivation to engage in a behavior has shown a positive impact on seeking information and sharing information [25]. In the context of health information, health motivators impact an individual's decisions to seek information [26].

When investigating behavioral intention, many studies adopt theories that focus on the assessment of costs and benefits such as the expectancy theory [27]. However, these factors are unstable traits and therefore measuring them does not offer health care providers constant estimation of their impact on patients' intentions. In general, patients have a positive attitude about the benefits of HIEs regarding improving health outcomes [15,17,28,29] and health motivators have an influence on people's intentions to engage in behaviors related to health outcomes [30]. Thus we have integrated patient activation into the model.

We define patient activation as the extent to which an individual wants to seek available resources and skills to engage in preventive health [31]. Patients' health lifestyles are relatively stable traits that provide constant estimation of people's intentions based on their health lifestyles measures [32]. Thus, patient activation is an appropriate stable predictor of patients' intentions to share their health information with providers.

> H_1: Patient activation is positively associated with the intention to share PHI electronically with providers.

Issue Involvement

Sharing PHI is influenced by personal contextual relevance [33], which is captured by issue involvement [6]. We define issue involvement as personal relevance, that is, how relevant a specific health issue is to each patient.

This study investigates the decision to share health information on the Web, and thus an involved person would be someone who is frequently visiting providers, has a number of diseases, and has a severe health status [6]. Issue involvement is defined in the literature as the extent to which an issue is of relevance and importance to the patient [34]. In the context of this study, the issue of concern is health and people's decisions to disclose health information electronically [6]. People who are more involved in the issue, health in this case, are expected to be more willing to share [33]. Therefore, the degree to which an individual may be contextually involved may influence his or her decision to share information. Thus, based on these results the second hypothesis is as follows:

> H_2: Issue involvement is positively associated with the intention to share PHI electronically.

Privacy

Privacy is defined as having control over who has access to medical records and being aware that those records are protected. With low control and no awareness of existing safeguards, the patients will have high privacy concerns.

A potential drawback in sharing information in HIE is the risk of patient privacy invasion and information security violations, which are increasing concerns due to the growing amount of health information exchanged electronically. Health care providers are driven by greater participation and systemic cost savings irrespective of benefits to individual patients. HIE is a new technology, and the risk of information breaches and privacy issues are not understood by the patients yet, especially when there is a lack of education. Moreover, when security breaches occur, patients are not compensated for their losses, which makes sharing of high privacy and security a risk. Simon et al [8] investigated the barriers to consent using 62 patients in a focus group and reported that privacy and security concerns, lack of knowledge of possible benefit to an individual's health, and the need for more information about the consent process are the main factors affecting sharing decisions.

While many people want to share their PHI with preferred providers [15,17], they are concerned with the privacy and security of their information with regard to HIEs.

> H_3: Privacy concern is negatively associated with the intention to share PHI electronically.

Trust

Trust is defined as the extent to which the patients have confidence in their health care providers. Lack of trust can be associated with other users who misuse or mishandle the information, with the system itself and its ability to protect information, or with people who illegally breach the system and misuse the information. The probability of any of these type or misuse happening increases when information is exposed to more people and shared and exchanged across multiple systems.

The more trust people have in the system the more engaged [35,36] they will be. For patients with HIV, trust in clinicians is associated with positive attitude toward sharing health information [14]. Therefore, a lack of trust is a barrier to patients consenting to share their PHI among providers. Trust has been associated with usability, that is, people will use a system more if they trust it. If patients are uncertain of why or how their information is used or shared, they can develop a lack of trust and thus be less engaging. Allowing patients to control who will access their health information and what information is available for access should lead to more engagement.

> H_4: Trust in providers is positively associated with the intention to share PHI electronically.

Patient-Physician Relationship

The time physicians spend with their patients is a dimension that defines the strength of the patient-physician relationship [37]. Communication and social factors such as warmth, feeling, help, and understanding are factors shaping a patient-physician relationship [38]. Therefore, a patient-physician relationship can be defined as the mixture of social strength, time spent, understanding, and help that characterizes the connection. When patients consent to share their PHI with providers, they are expecting physicians to use this information to make better decisions [17]. Thus, patients share information because they seek information represented in better doctors' opinions, diagnoses, and prescriptions. The strength of the patient-physician relationship is associated with patients' decisions to share information with their physicians [34,39].

> H_5: Patient-physician relationship is positively associated with the intention to share PHI electronically.

Methods

Data

A dataset from the National Cancer Institute was used to investigate the proposed hypotheses. The institute conducts the Health Information National Trend Survey known as HINTS, and this study used data collected in 2014 (HINTS 4 Cycle 4). The sample was nationally representative and thus made the findings generalizable.

As reported in the literature review, most studies have limitations for generalizability. The survey asks questions about participants' health conditions and health lifestyles, intention to share PHI, and a variety of related questions. Data from 1606 participants were used to analyze the conceptual model. Table 1 shows a distribution of participants. Of the total, 38.61% (620/1606) were males and 61.39% (986/1606) were females. The survey targeted adults who were aged 18 years and above. The average age was about 54 years, with a standard deviation of approximately 16. Among the participants, 18.99% (305/1606) have a high school degree, 60.33% (969/1606) have a college degree, and 20.67% (332/1606) have a postgraduate degree.

Measurement

Education, age, race, income level, health insurance, and the use of the Internet to look for health-related information were used as controls in the model. A formative measure of issue involvement was constructed using three items: the number of chronic diseases, the frequency of doctors' visits, and health status [6]. The factor loading for each of the items were 0.81, 0.76, and 0.67 as indicated in Multimedia Appendix 1. A composite factor score was calculated for the variable and the new variable was used in structural equation modeling (SEM) as an observed variable. Table 2 shows the results of the factor analysis. Since the factor analysis was conducted based on correlations rather than covariance, the use of different scales for each item was not an issue. The rest of the measures were constructed using reflective items. Intention to share PHI electronically consisted of six items, patient-physician relationship consisted of seven items, patient activation consisted of four items, and privacy concern involved two items. Trust in providers consisted of a single-item variable. Single items were acceptable if the question did not leave room for interpretation [39]. Single-item variables were used in information systems research that used SEM in the health care context [6]. The constructs were validated using confirmatory factor analysis and exploratory factor analysis (EFA). Table 2 shows the items and scales for all independent variables.

Reliability and Validity

We used STATA version 14.1 (StataCorp LP, College Station) to recode and analyze the data. To assess the validity of our measures, we performed confirmatory factor analyses on all questionnaire items using the STATA 14 SEM tool. The results reported in Table 2 show that all factor loadings for the constructs are substantial and significant. The goodness of fit indices—comparative fit index (CFI), root mean square error of approximation (RMSEA), and standardized root mean square residual (SRMR)—indicated a good fit of the measurements model for both pre- and post-intervention [40]. According to Hu and Bentler [40], the threshold for CFI is >0.90, for RMSEA is <0.06, and for SRMR is <0.10. The threshold for the Tucker-Lewis index (TLI), which is also known as the non-normed fit index, is >0.90. The Cronbach alpha values were all well above the threshold value of point 0.70 [41]. The composite reliability scores were also all well above the threshold of 0.70 [42]. Table 2 shows the Cronbach alpha and composite reliability scores for the measures. Convergent validity was assessed by calculating the average variance extracted (AVE), where each indicator was related to only one construct. The AVE values for all constructs exceeded 0.5, which was the desirable cutoff suggesting a convergent validity [43] (see Table 2). Discriminant validity was established for the study, because the AVE values for any two constructs exceeded the squared construct intercorrelation for each pair [43].

Table 1. Sample demographics (N=1606).

Demographic characteristics	Sample size, n (%)
Gender	
Male	620 (38.61)
Female	986 (61.39)
Age (in years)	
18–25	53 (3.30)
26–35	200 (12.45)
36–45	244 (15.19)
46–55	338 (21.05)
56–65	395 (24.6)
Over 65	376 (23.41)
Income (US $ **)**	
Under $10,000	97 (6.04)
$10,000 to under $15,000	87 (5.42)
$15,000 to under $20,000	86 (5.35)
$20,000 to under $35,000	189 (11.77)
$35,000 to under $50,000	234 (14.57)
$50,000 to under $75,000	300 (18.68)
$75,000 to under $100,000	230 (14.32)
$100,0000 or more	383 (23.85)
Education	
High school or less	305 (18.99)
Some college (2 years)	509 (31.69)
College degree (4 years)	460 (28.64)
Postgraduate degree	332 (20.67)
Number of chronic diseases	
0	489 (30.45)
1	453 (28.21)
2	309 (19.24)
3	220 (13.70)
4	85 (5.29)
5	42 (2.62)
6	8 (0.50)

Table 2. Factor analysis, reliability, and validity.

Construct items	CFA[a] factor loadings	EFA[b] factor loadings					Cronbach alpha	CR[c]	AVE[d]
		1	2	3	4	5			
Patient-physician relationship (PPR)							.94	0.9	0.67
PPR1	0.78	0.84	0.01	0.03	0	−0.03			
PPR2	0.80	0.83	0	0.02	−0.03	0.05			
PPR3	0.83	0.87	−0.01	−0.01	0.01	−0.02			
PPR4	0.87	0.89	−0.02	−0.01	−0.01	−0.02			
PPR5	0.82	0.86	−0.02	−0.03	0.02	−0.03			
PPR6	0.82	0.85	0.03	0.01	0.01	0.01			
PPR7	0.79	0.82	0.01	0	0	0.06			
Intention to share (INT)							.94	0.9	0.68
INT1	0.80	−0.02	0.86	−0.02	0	0.02			
INT2	0.84	0	0.89	−0.03	0.02	0			
INT3	0.87	−0.03	0.91	−0.02	0	0.02			
INT4	0.78	0.02	0.83	0.05	−0.01	−0.04			
INT5	0.85	0.03	0.88	0	0.02	−0.03			
INT6	0.81	0.01	0.86	0.02	−0.02	0.01			
Patient activation (PA)							.92	0.9	0.67
PA1	0.81	0.03	0	0.88	−0.02	−0.01			
PA2	0.75	−0.02	0	0.85	−0.04	0.09			
PA3	0.87	−0.01	0	0.90	0.04	−0.02			
PA4	0.84	0.01	0	0.89	0.03	−0.06			
Issue involvement (II)							.61	N/A	N/A
II1	N/A	−0.04	0	−0.03	0.81	−0.04			
II2		−0.05	0.03	0.08	0.67	0.05			
II3		0.1	−0.03	−0.05	0.76	0.01			
Privacy concern (PC)							.75	0.7	0.59
PC1	0.75	0	−0.04	−0.01	0.01	0.89			
PC2	0.78	0.03	0.04	0.01	0.01	0.88			
Overall goodness of fit									
RMSEA[e]	0.056								
CFI[f]	0.966								
TLI[g]	0.958								
SRMR[h]	0.025								

[a] CFA: confirmatory factor analysis.

[b] EFA: exploratory factor analysis.

[c] CR: composite reliability.

[d] AVE: average variance extracted.

[e] RMSEA: root mean square error of approximation.

[f] CFI: comparative fit index.

[g] TLI: Tucker-Lewis index.

[h] SRMR: standardized root mean square residual.

Table 3 shows the correlation matrix and the discriminant validity. The factor loadings of EFA showed strong loadings for items that belong to the same construct and very low loadings for items that belong to different constructs. No cross-loadings were observed, thus further establishing discriminant and convergent validity (see Table 2).

Common Method Variance

Data collected through a common method can suffer from common method variance (CMV), in which the relationship between the constructs is affected by the use of a single method [44]. CMV was assessed through a marker variable technique [45]. A marker variable is a variable that is theoretically unrelated to one or more of the variables measured in the study. *Worrying* was used as a marker variable that is theoretically unrelated to patient activation. The construct variables and the theoretically unrelated variable should have a low correlation. The correlation between the marker variable and health knowledge was 0.004 and not statistically significant, thus meeting the threshold of being below 0.1 [45]. After controlling for the marker variable using the approach introduced by Lindell and Whitney [45], the level of significance and the direction of the correlation between patient activation and every other variable did not change. Therefore, there is no evidence that the data was biased due to CMV.

Table 3. Correlation matrix and discriminant validity.

Variable	Average	SD	1	2	3	4	5	6	7	8	9
PPR[a]	1.66	0.68	0.819[e]								
Intention	2.8	0.92	0.045 (P=.07)	0.825[e]							
PA[b]	3.59	0.53	−0.043 (P=.08)	0.137 (P<.001)	0.819[e]						
II[c]	−0.04	0.99	0.077 (P<.001)	−0.051 (P=.04)	−0.103 (P<.001)	N/A					
PC[d]	1.94	0.52	0.261 (P<.001)	−0.120 (P<.001)	−0.022 (P=.37)	0.0077 (P=.76)	0.768[e]				
Trust	3.38	0.72	−0.665 (P<.001)	−0.024 (P 34)	0.047 (P=.06)	−0.062 (P=.01)	−0.248 (P<.001)	N/A			
Seek Internet	0.7	0.46	0.059 (P=.02)	0.182 (P<.001)	0.079 (P=.002)	−0.246 (P<.001)	0.056 (P=.02)	−0.057 (P=.02)	N/A		
Health insurance	0.92	0.27	−0.042 (P=.09)	−0.007 (P=.79)	0.009 (P=.70)	−0.036 (P=.15)	−0.033 (P=.19)	0.069 (P<.001)	0.046 (P=.06)	N/A	
Male	0.39	0.49	0.005 (P=.84)	−0.029 (P=.24)	−0.012 (P=.65)	−0.017 (P=.50)	0.083 (P<.001)	0.045 (P=.07)	−0.044 (P=.08)	0.065 (P=.009)	N/A
Age	53.77	15.8	−0.068 (P=.006)	−0.211 (P<.001)	−0.094 (P<.001)	0.314 (P<.001)	0.001 (P=.96)	0.092 (P<.001)	−0.291 (P<.001)	0.106 (P<.001)	0.164 (P<.001)

[a] PPR: patient-physician relationship.

[b] PA: patient activation.

[c] II: issue involvement.

[d] PC: privacy concern.

[e] Bold numbers in diagonal are the square root of average variance extracted.

Results

STATA version 14.1 was used to analyze the data. SEM was used to test the hypothesized model. Estimates derived from the SEM analysis were used to test the research hypotheses. The overall goodness of fit statistics of the structural model indicated a good model fit (RMSEA = 0.053, CFI = 0.91, and RMSR=0.068). In the first hypothesis, we proposed a relationship between patient activation and the intention to share PHI electronically. The path coefficient is positive and significant (βHM = 0.102, P<.001; see Table 4), suggesting that higher patient activation yields higher intention to share PHI electronically, thus supporting hypothesis 1. In the second hypothesis, we argued that higher issue involvement (II) would yield greater intention to share PHI electronically. The results support the second hypothesis (βII = 0.093, P=.001; see Table 4). The third hypothesis states that privacy concern (PC) is negatively associated with the intention to share PHI electronically. The path coefficient is negative and significant (βPC = −0.160, P<.001; see Table 4), suggesting that higher privacy concern yields lower intention to share PHI electronically. Thus, hypothesis 3 is supported. In hypothesis 4, we proposed a positive relationship between trust in providers and the intention to share PHI electronically. The trust in provider coefficient is −0.003 and not significant (P=.99), indicating that trust in providers does not play a significant role in people’s intentions to share their PHI electronically with the providers. There is no evidence to support hypothesis 4. Hypothesis 5 states that better patient-physician relationship (PPR) will yield higher intention to share PHI electronically. The path coefficient is positive and significant (βPPR = 0.103, *P*=.003; see Table 4), which provides support to hypothesis 5.

Table 4. Model results—dependent variable: intention to share personal health information electronically.

Variables	Standardized coefficients	*P* values	95% CI
Issue involvement	0.093	.001	0.036 to 0.149
Patient activation	0.102	<.001	0.051 to 0.153
Privacy concern	−0.160	<.001	−0.224 to −0.095
Patient-physician relationship	0.103	.003	0.034 to 0.171
Trust in providers	−0.003	.99	−1.39 to 1.385
Controls			
Seek Internet information	0.105	<.001	0.054 to 0.157
Have health insurance	−0.004	.88	−0.054 to 0.046
Male	0.005	.84	−0.043 to 0.054
Other controls (age, income, and education)			
Overall goodness of fit			
RMSEA[a]	0.053		
CFI[b]	0.91		
TLI[c]	0.90		
SRMR[d]	0.068		

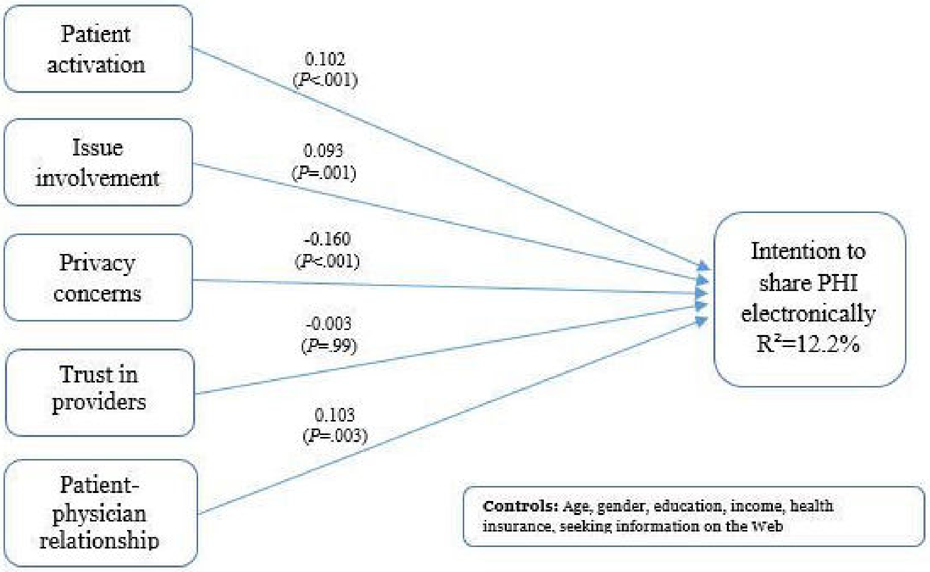
[a] RMSEA: root mean square error of approximation.

[b] CFI: comparative fit index.

[c] TLI: Tucker-Lewis index.

[d] SRMR: standardized root mean square residual.

Figure 2. Research model and results.



Discussion

Principal Findings

Overall, the results provide empirical support for the core hypotheses. Issue involvement, patient activation, privacy concern, and patient-physician relationship are all important and influenced individuals' intentions to share their PHI electronically with providers. We do not have data on actual behavior related to sharing PHI electronically with providers, but we suggest that this is an important area for future research. Trust in providers showed no significant relationship with the intention to share PHI. Among all the five main independent variables, privacy concern has the most influence on the intention to share. Standardized coefficients have been reported in Table 4. Standardized coefficient results show that the absolute magnitude for privacy concern is the highest with 0.160. The next highest number being patient-physician relationship is only 64.4% (0.103/0.160) of the magnitude of privacy concern. Given that privacy concern has a negative significant association unlike other significant variables, privacy concern is a major barrier to sharing PHI with providers. This finding provides practical implications to health care providers and policy makers of the significance of this concern. Health care providers and policy makers should prioritize their efforts and focus on addressing individuals' privacy concerns. In addition, health care providers should invest in educating people on the privacy policies that protect patients' information and privacy.

Historically, health care providers have focused on educating patients on the benefits of sharing their PHI, by emphasizing both cost and error reductions. Our study shows that there should be a shift in patient education, with a more salient focus on addressing privacy concerns. By making patients more aware of existing privacy policies and security measures in place, the health care providers are creating an environment where the patients are more likely to share their PHI, and therefore still able to achieve cost and error reduction benefits.

Patient-physician relationship has the highest positive significant magnitude. As patient-physician relationship consists of health professionals spending enough time with patients, involving patients in the decision making, helping patients understand steps needed to take care of their own health, and clarifying uncertainty, health professionals should pay particular attention to these factors.

Patient activation is a major factor associated with sharing behavior. It is also an essential element, because unlike trust, patient activation is a trait that is unlikely to change in the short run. Measuring patients' patient activation will provide outstanding insight into predicting sharing behavior.

Health care providers should focus more on people who are not involved in the issue because higher issue involvements will yield higher intention to share. Individuals who do not participate will be less likely to share, and thus health care providers should focus their attention in making the benefits of sharing more salient to noninvolved patients.

Overall, we propose a shift in education on two separate fronts: on the one hand, patient education is crucial to generate a perceived technological safe environment for sharing PHI electronically, and on another front, we suggest that physician education is as important as patient education. Physicians who are aware of the dimensions of the patient-physician relationship can improve the said relationship, leaving the patient more prone to PHI sharing, achieving better medical decisions, reduction in medical errors, and cost benefits.

Limitations

This study was based on a dataset that is publicly available. The constructs included in the study were therefore limited to those that are available in the collected survey. Another limitation is that the construct for measuring the trust in providers was based on a single-item variable. Wanous et al [39] state in their study, which specifically evaluates single-item measures, that single-item measures are reliable if there is little room for misunderstanding by the participants. For future studies, the trust in providers construct should be collected using multiple items.

Conclusions

This research contributes to the extended privacy calculus model. The privacy calculus model was used in this study because it is a very robust way to explain and predict people's intention to share their PHI. This assertion was substantiated by illustrating that the extended privacy calculus model is viable for explaining information sharing in the health care context.

In particular, this paper integrates personal interest variables in the health care context such as issue involvement and patient activation. The primary takeaway is that the research model provides theoretical and practical implications for sharing health care information. Privacy concerns are a central stage in modern society and the crown jewel is the sharing of PHI, and our model is a substantial first step in understanding the relevant variables related to sharing.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Operationalization of constructs.

[PDF File (Adobe PDF File), 23KB - jmir_v19i9e169_app1.pdf]

References

1. Frisse M, Johnson KB, Nian H, Davison CL, Gadd CS, Unertl KM, et al. The financial impact of health information exchange on emergency department care. J Am Med Inform Assoc 2012;19(3):328-333 [FREE Full text] [doi: 10.1136/amiajnl-2011-000394] [Medline: 22058169]
2. Frisse ME, Holmes RL. Estimated financial savings associated with health information exchange and ambulatory care referral. J Biomed Inform 2007 Dec;40(6 Suppl):S27-S32 [FREE Full text] [doi: 10.1016/j.jbi.2007.08.004] [Medline: 17942374]
3. Menachemi N, Collum TH. Benefits and drawbacks of electronic health record systems. Risk Manag Healthc Policy 2011;4:47-55 [FREE Full text] [doi: 10.2147/RMHP.S12985] [Medline: 22312227]
4. Yaraghi N, Du A, Sharman R, Gopal R, Ramesh R. Network effects in health information exchange growth. ACM Trans Manage Inf Syst 2013 Apr 01;4(1):1-26. [doi: 10.1145/2445560.2445561]
5. Clancy CM, Anderson KM, White PJ. Investing in health information infrastructure: can it help achieve health reform? Health Aff (Millwood) 2009;28(2):478-482 [FREE Full text] [doi: 10.1377/hlthaff.28.2.478] [Medline: 19276007]
6. Angst C, Agarwal R. Adoption of electronic health records in the presence of privacy concerns: the elaboration likelihood model and individual persuasion. MIS quarterly 2009;33(2):339-370.
7. Munoz RT, Fox MD, Gomez MR. Presumed consent models and health information exchanges: hard nudges and ambiguous benefits. Am J Bioeth 2013;13(6):14-15. [doi: 10.1080/15265161.2013.781718] [Medline: 23641837]
8. Simon SR, Evans J, Benjamin A, Delano D, Bates D. Patients' attitudes toward electronic health information exchange: qualitative study. J Med Internet Res 2009;11(3):e30 [FREE Full text] [doi: 10.2196/jmir.1164] [Medline: 19674960]
9. Caine K, Hanania R. Patients want granular privacy control over health information in electronic medical records. J Am Med Inform Assoc 2013 Jan 01;20(1):7-15 [FREE Full text] [doi: 10.1136/amiajnl-2012-001023] [Medline: 23184192]
10. Hassol A, Walker JM, Kidder D, Rokita K, Young D, Pierdon S, et al. Patient experiences and attitudes about access to a patient electronic health care record and linked web messaging. J Am Med Inform Assoc 2004;11(6):505-513 [FREE Full text] [doi: 10.1197/jamia.M1593] [Medline: 15299001]
11. Patel VN, Dhopeshwarkar RV, Edwards A, Barrón Y, Sparenborg J, Kaushal R. Consumer support for health information exchange and personal health records: a regional health information organization survey. J Med Syst 2010 Jul 29;36(3):1043-1052. [doi: 10.1007/s10916-010-9566-0]
12. Hersh W, Totten AM, Eden KB, Devine B, Gorman P, Kassakian SZ, et al. Outcomes from health information exchange: systematic review and future research needs. JMIR Med Inform 2015 Dec 15;3(4):e39 [FREE Full text] [doi: 10.2196/medinform.5215] [Medline: 26678413]
13. Ancker JS, Edwards AM, Miller MC, Kaushal R. Consumer perceptions of electronic health information exchange. Am J Prev Med 2012 Jul;43(1):76-80. [doi: 10.1016/j.amepre.2012.02.027] [Medline: 22704751]
14. Teixeira P, Gordon P, Camhi E, Bakken S. HIV patients' willingness to share personal health information electronically. Patient Educ Couns 2011 Aug;84(2):e9-12 [FREE Full text] [doi: 10.1016/j.pec.2010.07.013] [Medline: 20724095]
15. Dimitropoulos L, Patel V, Scheffler SA, Posnack S. Public attitudes toward health information exchange: perceived benefits and concerns. Am J Manag Care 2011 Dec;17(12 Spec No):SP111-SP116 [FREE Full text] [Medline: 22216769]
16. Grande D, Mitra N, Shah A, Wan F, Asch DA. Public preferences about secondary uses of electronic health information. JAMA Intern Med 2013 Oct 28;173(19):1798-1806 [FREE Full text] [doi: 10.1001/jamainternmed.2013.9166] [Medline: 23958803]
17. Patel VN, Dhopeshwarkar RV, Edwards A, Barron Y, Likourezos A, Burd L, et al. Low-income, ethnically diverse consumers' perspective on health information exchange and personal health records. Inform Health Soc Care 2011 Dec;36(4):233-252. [doi: 10.3109/17538157.2011.554930] [Medline: 21851182]
18. Li H, Gupta A, Zhang J, Sarathy R. Examining the decision to use standalone personal health record systems as a trust-enabled fair social contract. Decis Support Syst 2014 Jan;57:376-386. [doi: 10.1016/j.dss.2012.10.043]
19. Bansal G, Zahedi F, Gefen D. The impact of personal dispositions on information sensitivity, privacy concern and trust in disclosing health information online. Decis Support Syst 2010 May;49(2):138-150. [doi: 10.1016/j.dss.2010.01.010]
20. Dinev T, Hart P. An extended privacy calculus model for e-commerce transactions. Inf Syst Res 2006 Mar;17(1):61-80. [doi: 10.1287/isre.1060.0080]
21. Ajzen I. From intentions to actions: a theory of planned behavior. In: Khul J, Beckmann J, editors. Action Control. Philadelphia: Springer; 1985:11-39.
22. Ajzen I, Fishbein M. Understanding attitudes and predicting social behaviour. Upper Saddle River, NJ: Prentice-Hall; 1980.
23. Grant AM. Relational job design and the motivation to make a prosocial difference. Acad Manage Rev 2007 Apr 01;32(2):393-417. [doi: 10.5465/AMR.2007.24351328]
24. Grant AM. Does intrinsic motivation fuel the prosocial fire? motivational synergy in predicting persistence, performance, and productivity. J Appl Psychol 2008 Jan;93(1):48-58. [doi: 10.1037/0021-9010.93.1.48] [Medline: 18211134]
25. Reinholt M, Pedersen T, Foss N. Why a central network position isn't enough: the role of motivation and ability for knowledge sharing in employee networks. Acad Manag J 2011 Dec 01;54(6):1277-1297. [doi: 10.5465/amj.2009.0007]
26. Xiao N, Sharman R, Rao H, Upadhyaya S. Factors influencing online health information search: an empirical analysis of a national cancer-related survey. Decis Support Syst 2014 Jan;57:417-427. [doi: 10.1016/j.dss.2012.10.047]

27. Van Eerde W, Thierry H. Vroom's expectancy models and work-related criteria: a meta-analysis. J Appl Psychol 1996;81(5):575-586. [doi: 10.1037/0021-9010.81.5.575]
28. O'Donnell HC, Patel V, Kern LM, Barrón Y, Teixeira P, Dhopeshwarkar R, et al. Healthcare consumers' attitudes towards physician and personal use of health information exchange. J Gen Intern Med 2011 Sep;26(9):1019-1026 [FREE Full text] [doi: 10.1007/s11606-011-1733-6] [Medline: 21584839]
29. 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]
30. Fletcher S, Morgan TM, O'Malley MS, Earp JA, Degnan D. Is breast self-examination predicted by knowledge, attitudes, beliefs, or sociodemographic characteristics? Am J Prev Med 1989;5(4):207-215. [Medline: 2765291]
31. Jayanti R, Burns A. the antecedents of preventive health care behavior: an empirical study. J Acad Market Sci 1998 Jan 01;26(1):6-15. [doi: 10.1177/0092070398261002]
32. Moorman C, Matulich E. A model of consumers' preventive health behaviors: the role of health motivation and health ability. J Consum Res 1993 Sep;20(2):208-228. [doi: 10.1086/209344]
33. Loewenstein G. Hot-cold empathy gaps and medical decision making. Health Psychol 2005 Jul;24(4S):S49-S56. [doi: 10.1037/0278-6133.24.4.S49] [Medline: 16045419]
34. Cooper-Patrick L, Gallo J, Gonzales J. Race, gender, and partnership in the patient-physician relationship. JAMA 1999 Aug 11;282(6):583-589. [doi: 10.1001/jama.282.6.583]
35. Pavlou P, Fygensen M. Understanding and predicting electronic commerce adoption: an extension of the theory of planned behavior. MIS Quarterly 2006;30(1):115-143.
36. Williams H, Spencer K, Sanders C, Lund D, Whitley EA, Kaye J, et al. Dynamic consent: a possible solution to improve patient confidence and trust in how electronic patient records are used in medical research. JMIR Med Inform 2015 Jan 13;3(1):e3 [FREE Full text] [doi: 10.2196/medinform.3525] [Medline: 25586934]
37. Dugdale DC, Epstein R, Pantilat SZ. Time and the patient-physician relationship. J Gen Intern Med 1999 Jan;14(S1):S34-S40. [doi: 10.1046/j.1525-1497.1999.00263.x]
38. Larson EB, Yao X. Clinical empathy as emotional labor in the patient-physician relationship. JAMA 2005 Mar 02;293(9):1100-1106. [doi: 10.1001/jama.293.9.1100]
39. Wanous JP, Reichers AE, Hudy MJ. Overall job satisfaction: how good are single-item measures? J Appl Psychol 1997 Apr;82(2):247-252. [Medline: 9109282]
40. Hu L, Bentler P. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Struct Equ Modeling 1999 Jan;6(1):1-55. [doi: 10.1080/10705519909540118]
41. Nunnally J, Bernsein I. The assessment of reliability. Psychometric theory 1994;3(1):248-292.
42. Joreskog K, Sorbom D. LISREL 7: User reference guide. Chicago: Scientific Software; 1989.
43. Fornell C, Larcker DF. Structural equation models with unobservable variables and measurement error: algebra and statistics. J Mark Res 1981 Aug;18(3):382-388. [doi: 10.2307/3150980]
44. Podsakoff P, MacKenzie S, Lee JY, Podsakoff N. Common method biases in behavioral research: a critical review of the literature and recommended remedies. J Appl Psychol 2003;88(5):879-903. [doi: 10.1037/0021-9010.88.5.879]
45. Lindell MK, Whitney DJ. Accounting for common method variance in cross-sectional research designs. J Appl Psychol 2001 Feb;86(1):114-121. [Medline: 11302223]

Abbreviations

AVE: average variance extracted
CFA: confirmatory factor analyses
CFI: comparative fit index
CMV: common method variance
CR: composite reliability
EFA: exploratory factor analysis
HIE: health information exchange
HINTS: health information national trend survey
HITECH: health information technology for economic and clinicians health
II: issue involvement
PA: patient activation
PC: privacy concern
PHI: personal health information
PPR: patient-physician relationship
RMSEA: root mean square error of approximation
SEM: structural equation modeling
SRMR: standardized root mean square residual

TLI: Tucker-Lewis index

Edited by G Eysenbach; submitted 24.10.16; peer-reviewed by D Grande, A Gropper; comments to author 15.01.17; revised version received 03.03.17; accepted 18.03.17; published 13.09.17.

Please cite as:
Abdelhamid M, Gaia J, Sanders GL
Putting the Focus Back on the Patient: How Privacy Concerns Affect Personal Health Information Sharing Intentions
J Med Internet Res 2017;19(9):e169
URL: http://www.jmir.org/2017/9/e169/
doi:10.2196/jmir.6877
PMID:28903895

Original Paper

Reliability of an e-PRO Tool of EORTC QLQ-C30 for Measurement of Health-Related Quality of Life in Patients With Breast Cancer: Prospective Randomized Trial

Markus Wallwiener[1]; Lina Matthies[1]; Elisabeth Simoes[2]; Lucia Keilmann[1]; Andreas D Hartkopf[3]; Alexander N Sokolov[2]; Christina B Walter[3]; Nina Sickenberger[1]; Stephanie Wallwiener[1]; Manuel Feisst[4]; Paul Gass[5]; Peter A Fasching[5]; Michael P Lux[5]; Diethelm Wallwiener[3]; Florin-Andrei Taran[3]; Joachim Rom[1]; Andreas Schneeweiss[1]; Joachim Graf[2*]; Sara Y Brucker[2*]

[1]Gynecologic Oncology, National Center for Tumor Diseases, Hospital for General Obstetrics and Gynecology, University Hospital Heidelberg, Heidelberg, Germany

[2]Research Institute for Women's Health, Department of Women's Health, University Hospital Tuebingen, Tuebingen, Germany

[3]Department of Women's Health, University Hospital Tuebingen, Tuebingen, Germany

[4]Institute for Medical Biometry and Informatics, University Hospital Heidelberg, Heidelberg, Germany

[5]Department of Gynecology and Obstetrics, University Hospital Erlangen, Erlangen, Germany

[*]these authors contributed equally

Corresponding Author:
Markus Wallwiener
Gynecologic Oncology, National Center for Tumor Diseases, Hospital for General Obstetrics and Gynecology
University Hospital Heidelberg
Im Neuenheimer Feld 440
Heidelberg,
Germany
Phone: 49 (0)6221563
Email: markus.wallwiener@gmail.com

Abstract

Background: Breast cancer represents the most common malignant disease in women worldwide. As currently systematic palliative treatment only has a limited effect on survival rates, the concept of health-related quality of life (HRQoL) is gaining more and more importance in the therapy setting of metastatic breast cancer. One of the major patient-reported outcomes (PROs) for measuring HRQoL in patients with breast cancer is provided by the European Organization for Research and Treatment of Cancer (EORTC). Currently, paper-based surveys still predominate, as only a few reliable and validated electronic-based questionnaires are available. Facing the possibilities associated with evolving digitalization in medicine, validation of electronic versions of well-established PRO is essential in order to contribute to comprehensive and holistic oncological care and to ensure high quality in cancer research.

Objective: The aim of this study was to analyze the reliability of a tablet-based measuring application for EORTC QLQ-C30 in German language in patients with adjuvant and (curative) metastatic breast cancer.

Methods: Paper- and tablet-based questionnaires were completed by a total of 106 female patients with adjuvant and metastatic breast cancer recruited as part of the e-PROCOM study. All patients were required to complete the electronic- (e-PRO) and paper-based versions of the HRQoL EORTC QLQ-C30 questionnaire. A frequency analysis was performed to determine descriptive sociodemographic characteristics. Both dimensions of reliability (parallel forms reliability [Wilcoxon test] and test of internal consistency [Spearman rho and agreement rates for single items, Pearson correlation and Kendall tau for each scale]) were analyzed.

Results: High correlations were shown for both dimensions of reliability (parallel forms reliability and internal consistency) in the patient's response behavior between paper- and electronic-based questionnaires. Regarding the test of parallel forms reliability, no significant differences were found in 27 of 30 single items and in 14 of 15 scales, whereas a statistically significant correlation in the test of consistency was found in all 30 single items and all 15 scales.

Conclusions: The evaluated e-PRO version of the EORTC QLQ-C30 is reliable for patients with both adjuvant and metastatic breast cancer, showing a high correlation in almost all questions (and in many scales). Thus, we conclude that the validated paper-based PRO assessment and the e-PRO tool are equally valid. However, the reliability should also be analyzed in other prospective trials to ensure that usability is reliable in all patient groups.

Trial Registration: ClinicalTrials.gov NCT03132506; https://clinicaltrials.gov/ct2/show/NCT03132506 (Archived by WebCite at http://www.webcitation.org/6tRcgQuou).

(J Med Internet Res 2017;19(9):e322) doi:10.2196/jmir.8210

KEYWORDS

breast cancer; patient-reported outcomes; HRQoL; EORTC QLQ-C30; reliability

Introduction

Epidemiological Relevance of Breast Cancer

Breast cancer represents the most common malignant disease in women worldwide, with more than 71,000 new cases diagnosed every year in Germany [1]. In spite of improvement in progression-free survival (PFS) through promising targeted therapy options for hormone receptor-positive breast cancer, metastatic breast cancer remains an incurable disease [2-7]. In the year 2012, 17,853 women died of breast cancer in Germany alone [1]. Although early-stage breast cancer is often associated with high survival rates, the prognosis of metastatic breast cancer is significantly poorer, and therefore, the aim of treatment is mostly palliative because of minor probability of curation in patients with metastatic breast cancer [8]. Depending on the phenotype, median overall survival (OAS) after diagnosis of metastatic breast cancer is 2-3 years [9], ranging from 13.3 months for triple-negative [10] and 34.4 months for HER2-positive subtype of breast cancer [11].

Health-Related Quality of Life in Metastatic Breast Cancer

As currently systematic palliative treatment had only limited effect on survival rates, the concept of health-related quality of life (HRQoL) is gaining more and more importance in the therapy setting of metastatic breast cancer. Especially against the background of emerging side effects accompanying multiple oncological treatment lines, treatment should primarily aim at restoration and conservation of patients' HRQoL before prolonging the survival of patients [12]. Additionally, the diagnosis of an incurable disease represents an enormous emotional burden resulting in psychosocial distress that might impair the patient's well-being [1,13,14]. This is also taken into account by the recent German S3-guidelines for diagnosis, treatment, and aftercare of patients with breast cancer, recommending regular assessment of HRQoL during treatment [15]. HRQoL is defined as an *individual's perception of their position in life in the context of the culture and value systems in which they live and in relation to their goals, expectations, standards, and concerns* by World Health Organization [16]. HRQoL is divided into domains (1) health, (2) subjective feelings, (3) leisure time activities, (4) social relationships, (5) general activities, and (6) life satisfaction [17]. For many years, HRQoL in patients with breast cancer has been investigated. Especially in patients with metastatic disease, the measurement of HRQoL is important, as the primary goal of therapy is to afford them a high quality of life during their remaining lifespan [18-20].

Health-Related Quality of Life and Patient-Reported Outcomes

The most important obstacle has been the absence of widely accepted, standardized methods for carrying out such assessments, as much of the data suggest that clinicians miss or underestimate a large proportion of the symptomatic adverse events experienced by patients [21,22]. Furthermore, the assessment of adverse events and HRQoL by health care professionals is inconsistent when compared with the opinion of other professionals [23-25]. In this context, an independent report by the patient herself through patient-reported outcomes (PROs) could be more reliable and feasible. The US Food and Drug Administration (FDA) defines a PRO as *any report of the status of a patient's health condition that comes directly from the patient (or in some cases, a caregiver or surrogate), without interpretation of the patient's response by a clinician or anyone else* [26]. PROs comprise various aspects of the subjectively perceived state of health from patient's point of view, such as HRQoL, satisfaction with care, and drug adherence [27-29]. In this closely related areas, validated PROs are already the accepted gold standard for data collection, being used in clinical trials, as well as regulatory drug approvals (approximately 25% of US drug labels now include PRO-derived data) [30-33]. This aspect is also designated for optimization within the National Cancer Plan to enable information exchange between in-patient and outpatient treatment [30]. Concerning feasibility of patient's self-reported state of health in the oncological setting, a previous study demonstrated that most patients are willing and able to self-report their experiences with treatment [34]. Moreover, the Eastern Cooperative Oncology Group could not find any association between the patient's performance or functional status and compliance rates [30,35]. A study indicated that integrating HRQoL reports in daily clinical routine might represent an effective and time-saving option to improve medical care, as communication between patient and physician can be facilitated without extended interviews with the patient [36]. However, although regular assessment of HRQoL is generally recommended, routine evaluation is not yet provided in clinical practice [15,37]. One of the major PROs for measurement HRQoL in patients with breast cancer is provided by the European Organization for Research and Treatment of Cancer (EORTC). The EORTC QLQ-C30 as a modular approach is available in more than 100 languages and is used to assess

HRQoL in patients with various cancers within the scope of clinical trials, as well as in daily routine [38].

Electronic Monitoring of Patient-Related Outcome

With the expansion of digital tools, assessment of PRO by using an electronic equipment (e-PRO), such as tablet computers, is becoming a promising and economically viable approach, as real-time HRQoL monitoring allows early detection of patients at risk, ongoing improvement of oncological treatment, and ensuring the patient's safety [12]. Current data show that monthly compliance referring to frequency and completion of questionnaires with home Web reporting was high in breast cancer, warranting strategies to enhance compliance with routine care settings [30]. Additionally, some studies have suggested high feasibility of an electronic patient self-report platform in oncological patients, with mean compliance rates ranging from 75% to 85%, high patient satisfaction, and good usability of systems even among the non-Web avid and elderly patients [39,40]. Velikova et al [36] demonstrated that routine electronic HRQoL assessment in patients with breast cancer could positively influence physician-patient communication, potentially improving emotional functioning and HRQoL. Additionally, there is evidence that the completion of the questionnaire itself may improve the patient's well-being, regardless of whether the results are fed back to physicians [36]. However, knowledge regarding patient acceptance, feasibility, and barriers remains limited, especially in relation to health status, socioeconomic aspects, and the influence of other variables on patient's response behavior [41-44]. Currently, paper-based surveys of PRO still predominate, as there are only a few reliable and validated e-PRO questionnaires. The paper-based versions are frequently assigned into a tablet-based format without verification of reliability. As the aspects that influence the patient's willingness to use e-PRO and their response behavior by using e-PRO remain unclear, this strategy can endanger significance of e-PRO surveys [41]. For instance, the EORTC QLQ-C30 questionnaire has been used worldwide [45], but only reliable paper-based versions of it, although e-PRO has become much more prevalent (and "user-friendly") [46]. Facing the possibilities that are coming along with the evolving digitalization in medicine, the validation of electronic versions of well-established PRO is essential in order to contribute to a comprehensive and holistic oncological care and to ensure high quality in cancer research.

Aims and Objectives

The aim of this study was to analyze the reliability of a tablet-based e–PRO-measuring application for EORTC QLQ-C30 in German in patients with adjuvant (curative) and metastatic breast cancer compared with the established paper-based version. It should be analyzed if the response behavior of patients with breast cancer is influenced by the type of answering the questionnaire (answering by using paper and pencil or tablet computer) in a statistically significant way. We wanted to know whether there are differences in response behavior between the validated paper-based PRO version of EORTC QLQ-C30 and a new e-PRO version. The other aim was to examine the feasibility of using an e-PRO version of the EORTC QLQ-C30 for the future tablet-based measurement of HRQoL in patients with metastatic and adjuvant breast cancer in clinical practice. To achieve the aims, the patients were asked to fill out both paper- and tablet-based EORTC QLQ-C30 questionnaires.

Methods

Sample and Study Design

From July 2015 to May 2016, paper- and tablet-based PRO questionnaires were completed by a total 106 female patients with adjuvant and metastatic breast cancer treated consecutively at the Department of Women's Health in Tuebingen, Germany, and the National Cancer Centre in Heidelberg, Germany. The patients were recruited as part of the e-patient-reported outcomes and compliance analysis (PROCOM) study. The aim of e-PROCOM was to evaluate the general patient acceptance and practicability of a Web-based application for a PRO questionnaire for patients with adjuvant or metastatic breast cancer. The patients were asked to participate to compare the response behavior in paper- and Web-based questionnaires for analyzing the reliability of the e-PRO versions of the EORTC QLQ-C30 questionnaires (version 3) [38]. Inclusion criteria of e-PROCOM were female gender, full legal age (18 years and older), adjuvant or metastatic breast cancer diagnosis, sufficient language skills in German, and signed declaration on consent forms. Exclusion criterion was participation in other studies to minimize the burden of questionnaires. The patients were asked to complete the questionnaire during an outpatient visit to the hospital under the supervision of an attending physician. The study was designed as a double-centered (Tuebingen and Heidelberg), two-armed, prospective randomized trial. All patients were required to fill out the electronic- (e-PRO) and paper-based HRQoL questionnaire of EORTC QLQ-C30. Patients in arm A were assigned tablet computer, followed by paper questionnaire in the same session. Patients in arm B filled out the paper-based version, followed by the tablet-based questionnaire. The randomization procedure is based on the permuted-block randomization, which strives to generate equally large groups of treatment [47,48]. The postexposure acceptance for using the e-PRO tool was high (92%), as the patients were asked whether they could potentially imagine using tablet-based tools before using e-PRO [42]. Patients were informed about the aims of the study and asked for their consent ex ante. The study was approved by the Ethics Committee at the University of Tuebingen (project number 089/2015B02).

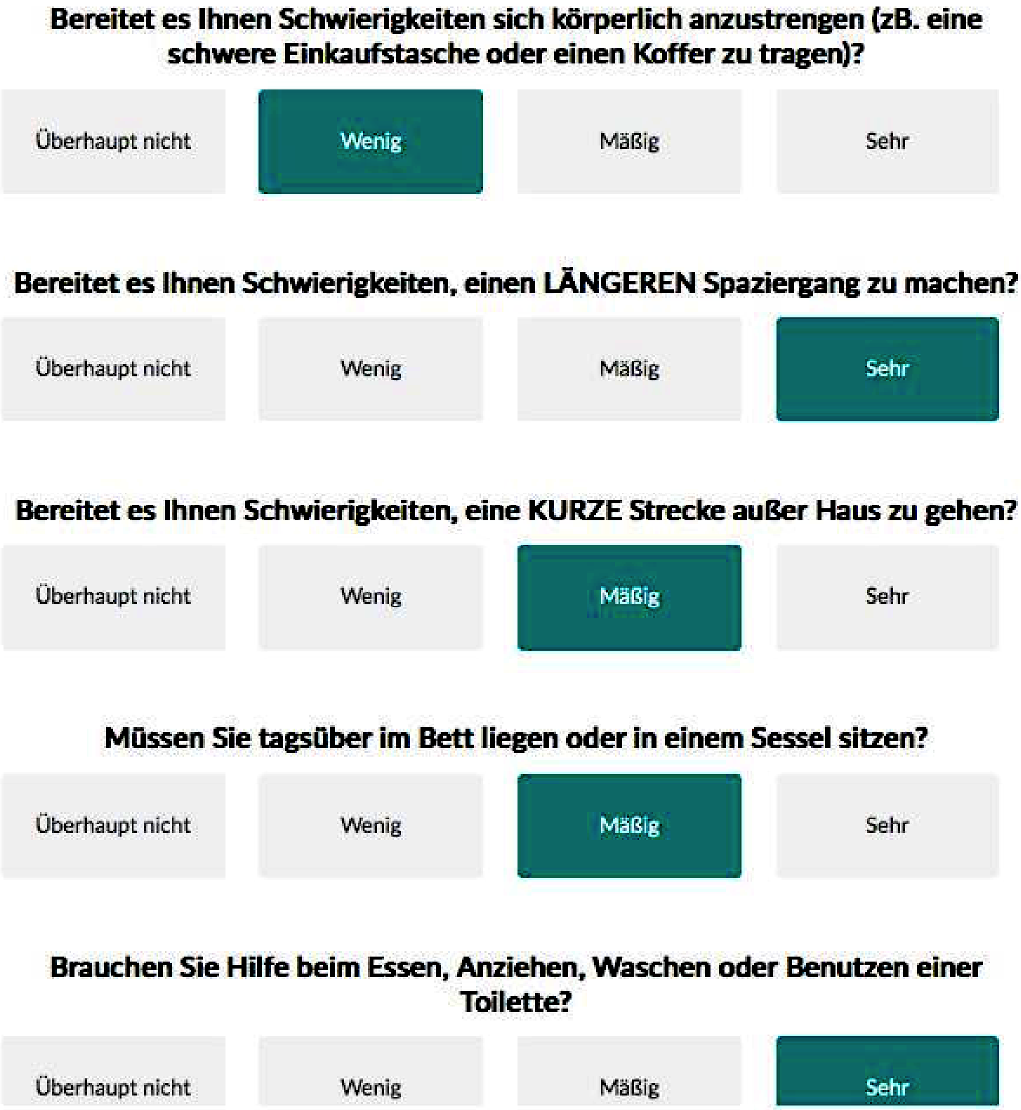
Procedure

The data collection was performed in 5 parts. The first part focused on the patients socio-economic variables. The second part contained the EORTC QLQ-C30, consisting of 30 questions in 5 subscales, various symptom scales, and individual items related to the patients health status on a multidimensional level. In addition, 28 of 30 questions are designed with a 4-point Likert scale and 2 questions with a 7-point Likert scale. Mean values were calculated in accordance with the official EORTC guidelines, which require a separate score to be calculated for each scale. The scores range from 0 to 100 [38,49]. The third part of the assessment also targeted HRQoL by administering FACT-B questionnaire, consisting of 37 questions with responses in 5 dimensions on a 5-point Likert scale [50,51]. The publication of FACT-B’s e-PRO reliability analysis is under preparation. The patients in the fourth part of the assessment were asked about preexisting technical skills, their willingness to use e-PRO, and potential barriers in relation to their health status [42], and the fifth part concerned the patients’ evaluation of the e-PRO tool (publication in preparation). The patients filled out the second and third parts of the assessment both paper- and tablet-based, whereas they answered only paper-based questionnaire in the other parts. In this study, we report the results of the second part of the assessment (reliability analyses of e-PRO tool of EORTC QLQ-C30).

Specifics of the PRO Tool

For e-PRO measurement, we used the PiiA (“Patient-informiert-interaktiv-Arzt”) Web-based application, which presents the relevant questions to be completed on a tablet computer. The PiiA-portal is a Web-based solution for capturing PROs, which was self-developed by the working group. Patients receive anonymous user credentials and are asked to complete FACT-B and QLQ-C30 questionnaires. Figure 1 shows the user interface of the first set of questions of the German EORTC QLQ-C30. The tool is constructed similar for all 28 questions with a 4-point Likert scale. Figure 2 shows the user interface of the 7-point Likert scale questions. After completing the questionnaires, patients log out and the pseudo-anonymized data will be backed up on a local storage device and securely locked.

Figure 1. Screenshot of PiiA application of the EORTC QLQ-C30 questionnaire for 4-point-scale questions (German).



Bereitet es Ihnen Schwierigkeiten sich körperlich anzustrengen (zB. eine schwere Einkaufstasche oder einen Koffer zu tragen)?

Überhaupt nicht | [✓ Wenig] | Mäßig | Sehr

Bereitet es Ihnen Schwierigkeiten, einen LÄNGEREN Spaziergang zu machen?

Überhaupt nicht | Wenig | Mäßig | [✓ Sehr]

Bereitet es Ihnen Schwierigkeiten, eine KURZE Strecke außer Haus zu gehen?

Überhaupt nicht | Wenig | [✓ Mäßig] | Sehr

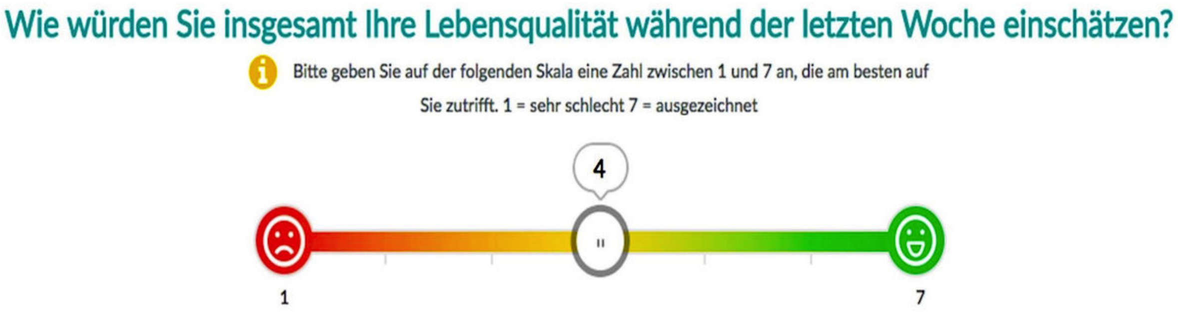
Müssen Sie tagsüber im Bett liegen oder in einem Sessel sitzen?

Überhaupt nicht | Wenig | [✓ Mäßig] | Sehr

Brauchen Sie Hilfe beim Essen, Anziehen, Waschen oder Benutzen einer Toilette?

Überhaupt nicht | Wenig | Mäßig | [✓ Sehr]

Figure 2. Screenshot of PiiA application of the EORTC QLQ-C30 questionnaire for 7-point-scalekaled questions (German).



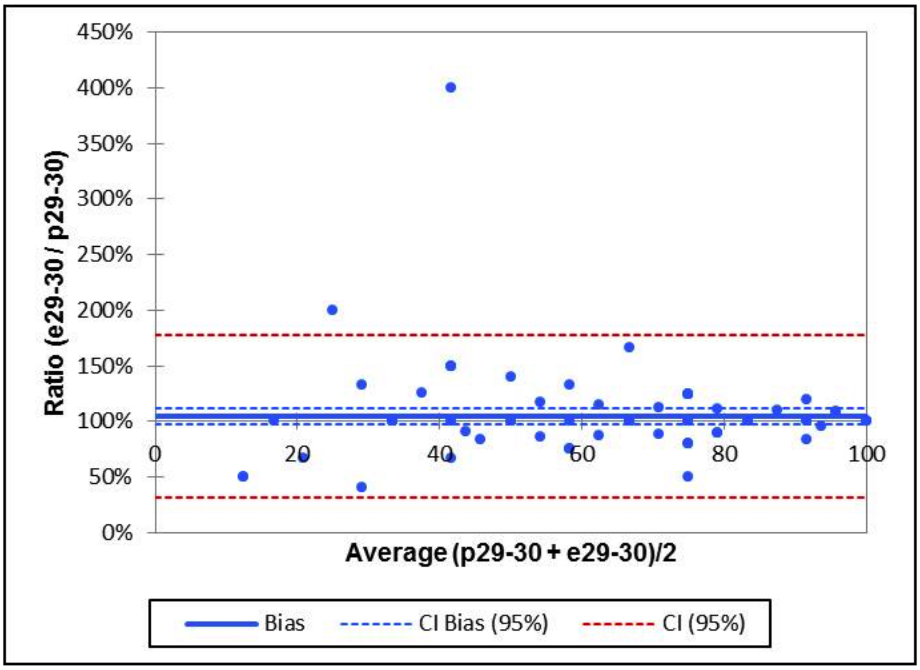
Wie würden Sie insgesamt Ihre Lebensqualität während der letzten Woche einschätzen?

Bitte geben Sie auf der folgenden Skala eine Zahl zwischen 1 und 7 an, die am besten auf Sie zutrifft. 1 = sehr schlecht 7 = ausgezeichnet

Statistical Analyses

A frequency analysis was first performed using IBM SPSS Statistics (version 24) to determine the descriptive sociodemographic characteristics of the patients. After that, we analyzed both dimensions of reliability (parallel forms reliability and test of internal consistency) and examined the disparity of responses and the rate of consistency between the paper-based PRO and e-PRO answers. Both dimensions of reliability were calculated for the 30 single items and 15 scales, resulting from the single items in accordance with the EORTC guidelines [49]. According to the Shapiro-Wilks test, the paired samples were not normally distributed, and therefore, we used the Wilcoxon test to identify possible statistically significant differences in the test of parallel forms reliability. Additionally, a Bland-Altman plot was created for the scale “overall state of health,” which represents the HRQoL (Figure 3). Earlier, the mean values of the paper-based PRO and the e-PRO measurements were calculated with the official EORTC guidelines [49], which require a separate score calculated for each scale, with scores ranging between 0 and 100. The consistency analyses were performed by calculation of correlation analyses (Spearman rho and agreement rates) for every EORTC question together with inter-item correlation (Pearson correlation) and rank correlation (Kendall tau) for each scale. With Pearson correlation, the internal consistency of a scale can be measured; it describes the extent to which the tasks or questions of a scale are interrelated. While Spearman rho test examines the internal consistency of the individual questions (specifically, the reliability of the e-PRO individual questions against the paper-based questions), Pearson correlation and Kendall tau tests are used to determine the scales calculated according to EORTC guidelines. Previously, we performed chi-square and Shapiro-Wilks tests between patients with metastatic and adjuvant breast cancer to identify possible statistically significant differences in relation to HRQoL. In all analyses, *P* values <.05 (2-tailed) were considered indicative of statistically significant differences (α=.05). As the analysis behaves as an explorative study, all reported *P* values can be received as purely descriptive. As we did not find any significant differences in relation to the response behavior between arms A and B in a pretest, we assessed both the arms together to compare the paper-based and e-PRO questionnaires of the patients. Bland-Altman plot was produced by using XLSTAT 2017. Missing values (which arose when patients did not answer individual questions) were ignored in the statistical calculation.

Figure 3. Bland-Altman plot for overall state of health.



Results

Patient Enrollment

Overall, 106 female patients with breast cancer were recruited, who completed both paper- and tablet-based EORTC QLQ-C30 questionnaires. Originally, n=153 patients were assessed for eligibility, of which 47 were excluded during recruiting, allocation, and data analyses as shown in the Consolidated Standards of Reporting Trials (CONSORT) flow diagram (Figure 4). A total of 53 patients were assigned tablet computer, followed by paper in the same session (arm A), whereas the same number of patients filled out the paper-based version, followed by the tablet-based questionnaire (arm B). Patients who had not completed more than half of the EORTC questions either paper- or tablet-based were excluded (1 patient in arm A and 2 patients in arm B). We did not find significant differences between the two arms in the response behavior likewise in sociodemographic status and in therapy setting, wherefore both arms were appreciated together. Previously, both arms were compared in all single items. Furthermore, 10 patients (arm A) and 16 patients (arm B), respectively, produced missing data in some questions (more often in the tablet-based questionnaire).

Sociodemographic Variables

Tables 1 and 2 show the sociodemographic characteristics of the study group, with 72% patients in adjuvant therapy and 28% in metastatic situation. We did not find significant intragroup differences between patients with metastatic and adjuvant breast cancer either in e-PRO or in paper-based PRO. Although the adjuvant and metastatic patients differ by focalizing age and HRQoL as the metastatic patients were older and reported a poorer HRQoL, we found no differences between the e-PRO response behavior of both groups. There were no differences in reliability of all single items and scales between metastatic and adjuvant patients because of which the whole collective was appreciated together. The mean age of the whole collective amounted to 51.0 years, and nearly one-third of the patients showed a higher level of education (high school diploma).

Figure 4. The CONSORT flow diagram.

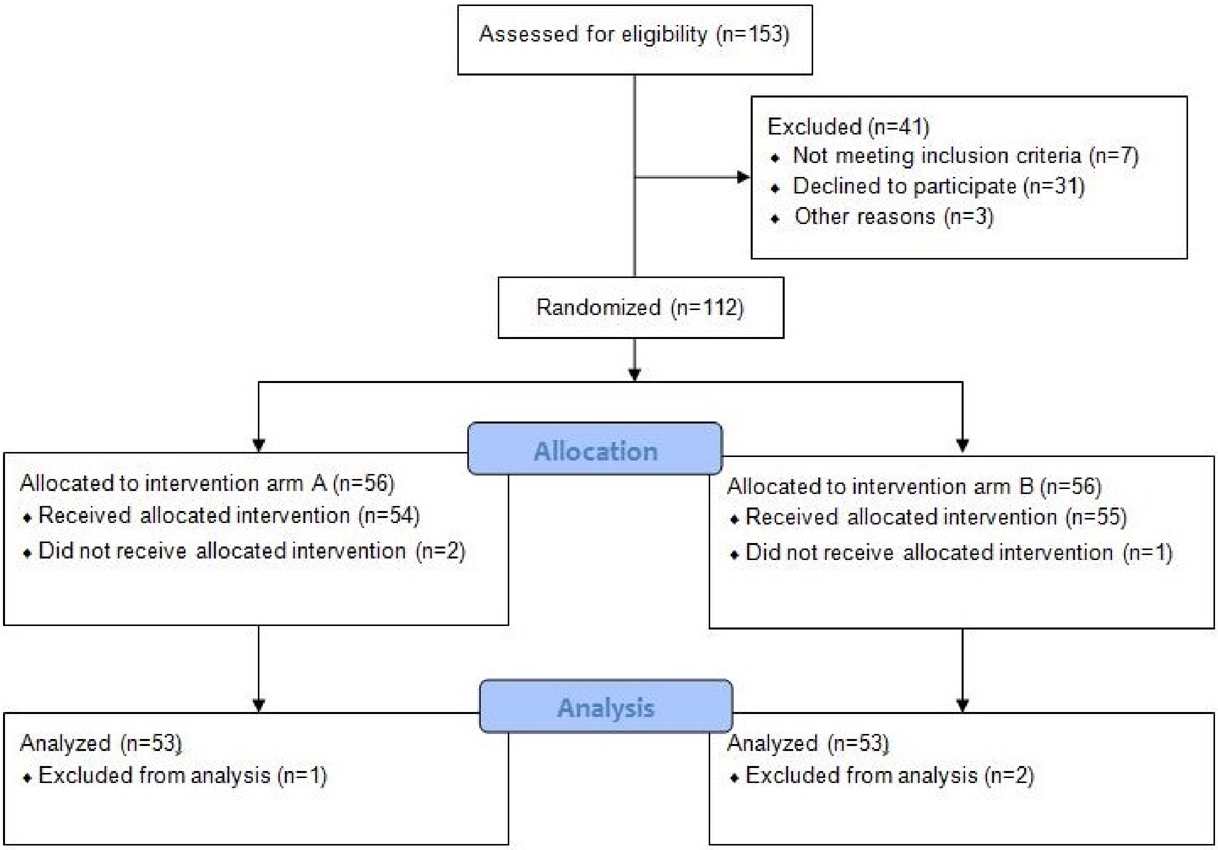


Table 1. Sociodemographic characteristics of the patients, part I.

Sociodemographic characteristics	Descriptive analyses (N=106)
Age	
Mean (SD)	51.0 (11.31)
Median (range, minimum-maximum)	52 (54, 30-84)
Level of education (1=lowest; 5=highest)	
Median	3.0
Interquartile range (25%-Quartil; 75%-Quartil)	2.0 (3.0; 5.0)

Table 2. Sociodemographic characteristics of the patients, part II.

Sociodemographic characteristics	n (%)	95% CI
Kind of education		
No qualification	1 (.9)	0-6
Main or secondary school leaving certificate	43 (40.6)	32-50
Advanced technical certificate	19 (17.9)	10-26
High school diploma (“Abitur”)	33 (31.1)	22-40
Not specified	10 (9.4)	2-15
Therapy setting		
Metastatic	30 (28.3)	19-35
Adjuvant treatment	76 (71.7)	61-83

Parallel Forms Reliability

The e-PRO tool seems to be reliable in the dimension of parallel forms reliability, as only few significant differences could be found. Table 3 shows the results of the Wilcoxon test of the 30 single items in the EORTC QLQ-C30. In only 3 items (in relation to *tiredness* and *pain* and *need to rest*), there were weak statistically significant differences between paper-based PRO and e-PRO. Tiredness was ranked a bit higher in the e-PRO questionnaire with noticeable differences by focalizing median’s characteristics ($MD_{\text{Paper-based PRO}}$=2.0 vs $MD_{\text{e-PRO}}$=3.0) similar to need to rest, whereas the patients confirmed more pain in the paper-based questionnaire. In 27 of the questions in the EORC QLQ-C30, there were no statistically significant differences between the patient’s answers in the paper-based questionnaire and e-PRO.

No statistically significant differences could be found in the Wilcoxon test for scoring values in the function and symptom scales of the EORTC QLQ-C30, except for the score “tiredness,” where the difference was significantly low (Table 4). The patient’s response behavior between paper-based PRO and e-PRO was similar on the 5 functional scales, the 9 symptom scales, and in the overall state of health, as the identified differences in the single items had balanced each other in relation to the scoring values’ calculation.

Figure 3 shows a Bland-Altman plot of the scale “overall of health” to compare agreement of measurements in a graphical way by plotting the difference between paper-based PRO and e-PRO against their mean. It is apparent that the deviations from the mean value of the difference are almost exclusively within the confidence interval, so the response behavior between paper-based PRO and e-PRO is not significantly different in the collective.

Table 3. Parallel forms reliability (Wilcoxon test) in single items.

Scale no.	Item no.	Short description of item	Paper-based PRO[a]		Electronic-based PRO		*P* value
			Mean (SD)	Median (interquartile range)	Mean (SD)	Median (interquartile range)	
I	1	Strenuous activities	2.40 (1.13)	2.0 (2.0)	2.33 (1.11)	2.0 (2.0)	.18
I	2	Long walk	2.19 (1.10)	2.0 (2.0)	2.23 (1.13)	2.0 (2.0)	.79
I	3	Short walk	1.56 (.85)	1.0 (1.0)	1.58 (.86)	1.0 (1.0)	>.99
I	4	Stay in bed or chair	1.79 (.87)	2.0 (1.0)	1.75 (.79)	2.0 (1.0)	.29
I	5	Self-care	1.18 (.55)	1.0 (.0)	1.14 (.51)	1.0 (0.0)	>.99
II	6	Limited in work	2.22 (1.05)	2.0 (2.0)	2.28 (.99)	2.0 (1.0)	.19
II	7	Hobbies or limited leisure activities	2.30 (1.10)	2.0 (2.0)	2.38 (1.04)	2.0 (1.0)	.18
III	21	Tense	2.06 (.85)	2.0 (2.0)	2.06 (.80)	2.0 (2.0)	.82
III	22	Worried	2.40 (.90)	2.0 (1.0)	2.49 (.81)	3.0 (1.0)	.31
III	23	Irritated	2.01 (.78)	2.0 (2.0)	2.02 (.79)	2.0 (2.0)	.71
III	24	Depressed	2.12 (.90)	2.0 (2.0)	2.18 (.91)	2.0 (2.0)	.85
IV	20	Concentration	1.86 (.97)	2.0 (1.0)	1.95 (1.06)	2.0 (2.0)	.06
IV	25	Memory difficulties	1.93 (.90)	2.0 (2.0)	1.96 (.91)	2.0 (2.0)	.53
V	26	Family life	2.21 (.96)	2.0 (2.0)	2.36 (1.01)	2.0 (1.0)	.09
V	27	Social life	2.39 (1.02)	2.0 (1.0)	2.41 (1.05)	2.0 (1.0)	.64
VI	10	Need to rest	2.33 (.93)	2.0 (1.0)	2.48 (.88)	2.0 (1.0)	*.005* [b]
VI	12	Felt weak	2.26 (.99)	2.0 (1.0)	2.20 (1.02)	2.0 (2.0)	.59
VI	18	Felt tired	2.48 (.93)	2.0 (1.0)	2.55 (.89)	3.0 (1.0)	*.03* [b]
VII	14	Nausea	1.54 (.78)	1.0 (1.0)	1.55 (.80)	1.0 (1.0)	.95
VII	15	Vomiting	1.09 (.45)	1.0 (.0)	1.13 (.54)	1.0 (.0)	>.99
VIII	9	Had pain	1.94 (.93)	2.0 (2.0)	1.92 (.85)	2.0 (2.0)	*.02* [b]
VIII	19	Pain interfered	1.95 (.99)	2.0 (2.0)	1.91 (.93)	2.0 (1.5)	.19
IX	8	Shortness of breath	1.92 (.87)	2.0 (1.0)	1.84 (.87)	2.0 (1.0)	.27
X	11	Sleep disturbance	2.15 (1.01)	2.0 (2.0)	2.23 (1.05)	2.0 (2.0)	.13
XI	13	Lack of appetite	1.56 (.88)	1.0 (1.0)	1.54 (.87)	1.0 (1.0)	.89
XII	16	Constipation	1.52 (.78)	1.0 (1.0)	1.47 (.70)	1.0 (2.0)	.95
XIII	17	Diarrhea	1.43 (.76)	1.0 (1.0)	1.40 (.72)	1.0 (1.0)	.15
XIV	28	Financial impact of disease	1.97 (1.03)	2.0 (2.0)	1.93 (1.03)	2.0 (2.0)	.28
XV	29	Physical condition	4.63 (1.43)	5.0 (2.0)	4.63 (1.37)	5.0 (2.0)	.95
XV	30	General QoL[c]	4.67 (1.49)	5.0 (2.13)	4.72 (1.59)	5.0 (2.0)	.29

[a] PRO: Patient reported outcomes.

[b] Statistically weak significant difference.

[c] QoL: quality of life.

Table 4. Parallel forms reliability (Wilcoxon test) for scoring values in the function and symptom scales.

Scale	Number of questions (items) QLQ-C30	Paper-based PRO[a]		Electronic-based PRO		*P* value
		Mean (SD)	Median (interquartile range)	Mean (SD)	Median (interquartile range)	
Functional scales						
Physical resilience[a] (I)	5 (1-5)	73.14 (24.83)	80.0 (33.3)	73.89 (23.65)	80.0 (33.3)	.08
Resilience at work and during leisure time activities[a] (II)	2 (6-7)	57.92 (34.61)	66.67 (50)	55.77 (32.41)	66.67 (50)	.59
Emotional resilience[a] (III)	4 (21-24)	61.53 (23.95)	66.67 (41.67)	60.88 (23.22)	58.33 (33.34)	.53
Cognitive resilience[a] (IV)	2 (20-25)	70.18 (26.00)	66.67 (50)	68.67 (27.67)	66.67 (50)	.12
Social resilience[a] (V)	2 (26,27)	56.93 (30.46)	66.67 (50)	54.23 (32.58)	66.67 (50)	.12
Symptom scales						
Tiredness (VI)	3 (10,12,18)	45.21 (28.41)	33.33 (36.11)	45.10 (28.08)	33.33 (44.44)	*.05* [b]
Nausea or vomiting (VII)	2 (14,15)	10.42 (18.13)	.0 (1.47)	11.52 (19.80)	.0 (3.19)	.28
Pain (VIII)	2 (9,19)	31.25 (30.46)	33.33 (50.0)	29.34 (26.65)	33.33 (50.0)	.35
Shortness of breath (IX)	1 (8)	30.58 (29.06)	33.33 (33.33)	28.00 (29.10)	33.33 (33.33)	.70
Sleep disturbance (X)	1 (11)	38.45 (33.64)	33.33 (66.67)	41.06 (34.84)	33.33 (66.67)	.71
Lack of appetite (XI)	1 (13)	18.65 (29.18)	.0 (33.33)	18.14 (29.04)	.0 (33.3)	.72
Constipation (XII)	1 (16)	17.31 (25.87)	.0 (33.33)	15.69 (23.37)	.0 (33.33)	.47
Diarrhea (XIII)	1 (17)	14.22 (25.41)	.0 (33.33)	13.27 (23.96)	.0 (33.33)	.19
Financial impact of disease (XIV)	1 (28)	32.19 (34.33)	33.33 (66.67)	31.05 (34.40)	33.33 (66.67)	.20
Overall state of health[a] (XV)	2 (29,30)	60.78 (23.75)	66.67 (41.67)	61.30 (23.82)	66.67 (33.33)	.52

[a] Items are scaled from worst to best, with high scores representing a good QoL profile.

[b] Statistically weak significant difference.

Test of Internal Consistency

Table 5 shows the Spearman rho correlation values and agreement rates, which were obtained for every question of the EORTC QLQ-C30 questionnaire. In all 30 questions, a high correlation (>.7) was found between paper-based PRO and e-PRO. In 23 questions, the correlation levels was >.85, and in 1 question, we found a maximal correlation of 1.0. In all 30 correlated questions, agreement rates fluctuated between 66.6% and 100%.

Table 6 shows the results of internal consistency testing for the function and the symptom scales of the EORTC between paper-based PRO and e-PRO. There were high correlations in the response behavior (>.7) in all 5 functional scales, all 9 symptom scales, and the overall state of health. Statistically high significant correlations (>.9) between paper-based PRO and e-PRO was found in all 5 functional scales by focalizing interitem correlation (Pearson correlation). The rank correlation in all functional scales was also high, as Kendall tau coefficient ranged between .79 and .92. In the overall state of health, the correlation was .88 (Pearson correlation) and .77 (Kendall tau). In the symptom scales, the consistency rates were between .80 (*shortness of breath*) and .96 (*lack of appetite*) in relation to interitem correlation, together with rates between .71 (*shortness of breath*) and .93 (*diarrhea*) with regard to rank correlation.

Table 5. Test of internal consistency in single items: results of correlation (Spearman rho) and agreement analysis.

Scale no.	Item no.	Short description of items	Spearman rho	*P* value of Spearman rho[b]	Agreement (%)
I	1	Strenuous activities	.932	<.001	81.4
I	2	Long walk	.949	<.001	86.6
I	3	Short walk	.930	<.001	91.8
I	4	Stay in bed or chair	.848	<.001	84.1
I	5	Self-care	.999	<.001	97.4
II	6	Limited in work	.863	<.001	76.8
II	7	Hobbies or limited leisure activities	.867	<.001	81.8
III	21	Tense	.936	<.001	74.3
III	22	Worried	.846	<.001	80.6
III	23	Irritated	.945	<.001	90.2
III	24	Depressed	.812	<.001	77.2
IV	20	Concentration	.900	<.001	79.5
IV	25	Memory difficulties	.936	<.001	86.2
V	26	Family life	.876	<.001	75.3
V	27	Social life	.891	<.001	78.9
VI	10	Need to rest	.860	<.001	78.2
VI	12	Felt weak	.878	<.001	85.3
VI	18	Felt tired	.829	<.001	78.8
VII	14	Nausea	.924	<.001	89.4
VII	15	Vomiting	1.00	<.001	100
VIII	9	Had pain	.774	<.001	78.9
VIII	19	Pain interfered	.895	<.001	81.0
IX	8	Shortness of breath	.792	<.001	77.3
X	11	Sleep disturbance	.909	<.001	81.9
XI	13	Lack of appetite	.947	<.001	92.4
XII	16	Constipation	.899	<.001	66.6
XIII	17	Diarrhea	.965	<.001	94.0
XIV	28	Financial impact of disease	.922	<.001	85.8
XV	29	Physical condition	.830	<.001	70.0
XV	30	General QoL[a]	.863	<.001	65.0

[a] QoL: quality of life.

[b] statistically high significant correlations.

Table 6. Test of internal consistency in the function scales and symptom scales: results of Pearson correlation and Kendall tau analyses.

Scale	Pearson correlation (95% CI)	*P* value[a]	Kendall tau	*P* value[a]
Functional scales				
Physical resilience	.979 (.966-.987)	<.001	.918	<.001
Resilience at work and during leisure time activities	.900 (.851-.933)	<.001	.795	<.001
Emotional resilience	.941 (.906-.963)	<.001	.848	<.001
Cognitive resilience	.914 (.865-.945)	<.001	.866	<.001
Social resilience	.921 (.874-.950)	<.001	.819	<.001
Symptom scales				
Tiredness	.948 (.916-.968)	<.001	.851	<.001
Nausea or vomiting	.956 (.928-.973)	<.001	.873	<.001
Pain	.907 (.852-.942)	<.001	.819	<.001
Shortness of breath	.798 (.712-.860)	<.001	.710	<.001
Sleep disturbance	.921 (.874-.951)	<.001	.810	<.001
Lack of appetite	.959 (.933-.974)	<.001	.916	<.001
Constipation	.869 (.793-.917)	<.001	.828	<.001
Diarrhea	.953 (.931-.968)	<.001	.928	<.001
Financial impact of disease	.919 (.872-.949)	<.001	.863	<.001
Overall state of health	.878 (.823-.916)	<.001	.769	<.001

[a] Statistically high significant correlations.

Discussion

Principal Findings

In both dimensions of reliability (parallel forms reliability and internal consistency), we found high correlations with only few differences in the patient's response behavior between paper-based PRO and e-PRO in the EORTC QLQ-C30 questionnaire. In the test of parallel forms reliability, we found statistically significant differences in only 3 of 30 questions. By focalizing the function scales and the symptom scales, there was only one statistically significant difference between the patient's answers in both procedures. In the dimension of consistency, there were high correlation and agreement rates in all items and scales. Due to only few differences and high correlations in almost all single items and scales, the PiiA tool's e-PRO version of EORTC QLQ-C30 seems to be reliable for HRQoL measurement in patients with metastatic and adjuvant breast cancer. Due to the results, we cannot expect that the future use of the PiiA tool in the same patient group will show significant differences between paper-based PRO and e-PRO version of QLQ-C30, or that patient' response behavior will be significantly influenced by the survey tool after transition to electronic-based patient surveys. Therefore, the tool is suitable for ascertaining HRQoL in patients with metastatic or adjuvant breast cancer.

Comparison With Prior Work

Although e-PRO applications are on the rise, paper-based surveys of PRO still predominate clinical research, as there is a lack of reliable electronically validated questionnaires. One of the most used questionnaires for the measurement of HRQoL, especially in patients with breast cancer, is the EORTC QLQ-C30, with reliable paper-based format in many languages but without reliable electronic-based version in German. Electronic-based utilization of EORTC QLQ-C30 and other PRO without verification of reliability could endanger significance of e-PRO surveys, wherefore corresponding analysis in relation to differences and correlations between the paper-based vision of EORTC QLQ-C30 and newly developed online tools is of great importance. It can be assumed that several aspects (ie, sociodemographics, technical skills, health condition, and maybe design specifics of the e-PRO tool) influence the patient's willingness to use e-PRO and their response behavior, which underlines the necessity of reliability analyses [41,42].

Limitations

However, there were some limitations in the study design and the methodological implementation, which could possibly reduce data's validity. In 3 questions of the test of parallel forms reliability, we found several missing values that could be because of the length of the survey. The patients were surveyed while they were receiving chemotherapy intervention, and they were not permitted to take the questionnaire home to complete it. Obviously, the length of the questionnaire had an effect on the patients' concentration, as missing values were found especially in the second response run of the EORTC questionnaire. Possibly, the burden of therapy was potentially affecting the ability of some patients to fill both paper- and tablet-based questionnaires during an outpatient visit. Missing values were found particularly in the questions "need to rest" and "felt tired," which were potentiated unfavorable to the dimension "tiredness" in the dimension of parallel forms reliability. In addition, the tumor stage, extent of metastasis, and the administered therapy were beyond the scope of study. Furthermore, psycho-oncological information was not collected, although psycho-oncological distress is a commonly associated burden that could potentially influence the willingness to use e-PRO and therefore e-PRO's reliability [42,52]. It was not possible to determine whether the state of health was lower and the psychological distress was higher in those patients who could not be motivated to participate in the study compared with those who could be included, wherefore a selection bias was in place. Therefore, it is not possible to assess conclusively whether the e-PRO version is reliable for all patients with metastatic and adjuvant breast cancer or only for those with substantiated willingness. It must also be ensured that the proven reliability for the mentioned patient group applies only to the electronic-based version of the EORTC QLQ-C30 questionnaire but not to the PiiA tool per se.

Strengths of the Study

The EORTC QLQ-C30 questionnaire has been used worldwide [45], but only with reliable paper-based versions, whereas e-PRO has become much more prevalent (and "user-friendly") [46]. It is one of the strengths of this study that a new tool for e-PRO measurement was developed and analyzed regarding reliability in patients with breast cancer, while other studies often assign paper-based versions into a tablet-based format without verification of reliability. Reliability was ascertained in a multidimensional way, as parallel forms reliability (Wilcoxon test) and internal consistency (by focalizing Spearman rho, agreement rates, Pearson correlation and Kendall tau) were calculated. Our data show that patients with both the adjuvant and metastatic breast cancer are able to use e-PRO, as the PiiA tool was reliable in both patient groups. However, willingness to use e-PRO in patients with metastatic and adjuvant breast cancer is dependent on technical exposition [42]. The results of the study can improve the quality of e-PRO measurements, as they seem to be generalizable, and the PiiA application of EORTC QLQ-C30 can be used for reliable e-based measurement of HRQoL in other studies and clinical routine. The tool is reliable in female patients with breast cancer, as hurdles for e-PRO could be found especially in metastatic patients [42].

Conclusions

Electronic-based PRO is constantly being adopted in clinical research and clinical routine, which underlines the necessity of reliable questionnaires. The evaluated PiiA's version of the EORTC QLQ-C30 is reliable for patients with breast cancer in adjuvant or metastatic setting because high correlation was found in almost all questions (and in many scales). Thus, we conclude equality between the validated p-PRO assessment and the used e-PRO tool. However, the reliability in other prospective trials should also be analyzed to ensure the reliable usability in all patient groups.

Conflicts of Interest

PG has received honoraria from Novartis and financial support for symposia from Novartis, Roche, and PharmaMar. The other authors declare that they have no conflict of interest.

Multimedia Appendix 1

CONSORT-EHEALTH checklist (v1.6.1).

[PDF File (Adobe PDF File), 1MB - jmir_v19i9e322_app1.pdf]

References

1. Krebsdaten. 2016. Bericht zum krebsgeschehen in Deutschland 2016 URL: http://www.krebsdaten.de/Krebs/DE/Content/Publikationen/Krebsgeschehen/Krebsgeschehen_download.pdf?__blob=publicationFile [accessed 2017-08-28] [WebCite Cache ID 6t3CWhJes]
2. Verma S, Miles D, Gianni L, Krop IE, Welslau M, Baselga J, et al. Trastuzumab emtansine for HER2-positive advanced breast cancer. N Engl J Med 2012 Nov 08;367(19):1783-1791 [FREE Full text] [doi: 10.1056/NEJMoa1209124] [Medline: 23020162]
3. Baselga J, Cortés J, Kim S, Im S, Hegg R, Im Y, et al. Pertuzumab plus trastuzumab plus docetaxel for metastatic breast cancer. N Engl J Med 2012 Jan 12;366(2):109-119. [doi: 10.1056/NEJMoa1113216] [Medline: 22149875]
4. Baselga J, Campone M, Piccart M, Burris HA, Rugo HS, Sahmoud T, et al. Everolimus in postmenopausal hormone-receptor-positive advanced breast cancer. N Engl J Med 2012 Feb 09;366(6):520-529. [doi: 10.1056/NEJMoa1109653] [Medline: 22149876]
5. Hortobagyi GN, Stemmer SM, Burris HA, Yap Y, Sonke GS, Paluch-Shimon S, et al. Ribociclib as first-line therapy for HR-positive, advanced breast cancer. N Engl J Med 2016 Dec 03;375(18):1738-1748. [doi: 10.1056/NEJMoa1609709] [Medline: 27717303]
6. Finn RS, Crown JP, Lang I, Boer K, Bondarenko IM, Kulyk SO, et al. The cyclin-dependent kinase 4/6 inhibitor palbociclib in combination with letrozole versus letrozole alone as first-line treatment of oestrogen receptor-positive, HER2-negative, advanced breast cancer (PALOMA-1/TRIO-18): a randomised phase 2 study. Lancet Oncol 2015 Jan;16(1):25-35. [doi: 10.1016/S1470-2045(14)71159-3] [Medline: 25524798]

7. Finn RS, Martin M, Rugo HS, Jones S, Im S, Gelmon K, et al. Palbociclib and letrozole in advanced breast cancer. N Engl J Med 2016 Dec 17;375(20):1925-1936. [doi: 10.1056/NEJMoa1607303] [Medline: 27959613]
8. Dufresne A, Pivot X, Tournigand C, Facchini T, Altweegg T, Chaigneau L, et al. Impact of chemotherapy beyond the first line in patients with metastatic breast cancer. Breast Cancer Res Treat 2008 Jan;107(2):275-279. [doi: 10.1007/s10549-007-9550-7] [Medline: 17380382]
9. Kucharczyk MJ, Parpia S, Walker-Dilks C, Banfield L, Swaminath A. Ablative therapies in metastatic breast cancer: a systematic review. Breast Cancer Res Treat 2017 Jul;164(1):13-25. [doi: 10.1007/s10549-017-4228-2] [Medline: 28401364]
10. Kassam F, Enright K, Dent R, Dranitsaris G, Myers J, Flynn C, et al. Survival outcomes for patients with metastatic triple-negative breast cancer: implications for clinical practice and trial design. Clin Breast Cancer 2009 Feb;9(1):29-33. [doi: 10.3816/CBC.2009.n.005] [Medline: 19299237]
11. Savci-Heijink CD, Halfwerk H, Hooijer GK, Horlings HM, Wesseling J, van de Vijver MJ. Retrospective analysis of metastatic behaviour of breast cancer subtypes. Breast Cancer Res Treat 2015 Apr;150(3):547-557 [FREE Full text] [doi: 10.1007/s10549-015-3352-0] [Medline: 25820592]
12. Wintner LM, Giesinger JM, Zabernigg A, Sztankay M, Meraner V, Pall G, et al. Quality of life during chemotherapy in lung cancer patients: results across different treatment lines. Br J Cancer 2013 Oct 29;109(9):2301-2308 [FREE Full text] [doi: 10.1038/bjc.2013.585] [Medline: 24091620]
13. Trautmann F, Hentschel L, Hornemann B, Rentsch A, Baumann M, Ehninger G, et al. Electronic real-time assessment of patient-reported outcomes in routine care-first findings and experiences from the implementation in a comprehensive cancer center. Support Care Cancer 2016 Jul;24(7):3047-3056. [doi: 10.1007/s00520-016-3127-0] [Medline: 26887586]
14. Goodwin PJ, Ennis M, Bordeleau LJ, Pritchard KI, Trudeau ME, Koo J, et al. Health-related quality of life and psychosocial status in breast cancer prognosis: analysis of multiple variables. J Clin Oncol 2004 Oct 15;22(20):4184-4192. [doi: 10.1200/JCO.2004.12.091] [Medline: 15483029]
15. Awmf. 2012. Interdisziplinaere S3-leitlinie fuer die diagnostik, therapie und nachsorge des mammakarzinoms URL: http://www.awmf.org/uploads/tx_szleitlinien/032-045OL_k_S3__Brustkrebs_Mammakarzinom_Diagnostik_Therapie_Nachsorge_2012-07.pdf [accessed 2017-08-28] [WebCite Cache ID 6t3Ck0VAW]
16. WHO. WHOQOL: measuring quality of life URL: http://www.who.int/healthinfo/survey/whoqol-qualityoflife/en/ [accessed 2017-08-28] [WebCite Cache ID 6t3CpL5TX]
17. Ritsner M, Modai I, Endicott J, Rivkin O, Nechamkin Y, Barak P, et al. Differences in quality of life domains and psychopathologic and psychosocial factors in psychiatric patients. J Clin Psychiatry 2000 Nov;61(11):880-9; quiz 890. [Medline: 11105747]
18. Cortés J, Baselga J, Im Y, Im S, Pivot X, Ross G, et al. Health-related quality-of-life assessment in CLEOPATRA, a phase III study combining pertuzumab with trastuzumab and docetaxel in metastatic breast cancer. Ann Oncol 2013 Oct;24(10):2630-2635. [doi: 10.1093/annonc/mdt274] [Medline: 23868905]
19. Korpela J, Mali P, Kaljonen A, Salminen E. Quality of life of patients with metastatic breast cancer treated with epirubicin and docetaxel. IJCM 2011;02(03):346-351. [doi: 10.4236/ijcm.2011.23060]
20. Karamouzis MV, Ioannidis G, Rigatos G. Quality of life in metastatic breast cancer patients under chemotherapy or supportive care: a single-institution comparative study. Eur J Cancer Care (Engl) 2007 Sep;16(5):433-438. [doi: 10.1111/j.1365-2354.2006.00771.x] [Medline: 17760930]
21. Raphael J, Verma S. Overall survival (OS) endpoint: an incomplete evaluation of metastatic breast cancer (MBC) treatment outcome. Breast Cancer Res Treat 2015 Apr;150(3):473-478. [doi: 10.1007/s10549-015-3342-2] [Medline: 25783185]
22. Stefanovic S, Wallwiener M, Karic U, Domschke C, Katic L, Taran F, et al. Patient-reported outcomes (PRO) focused on adverse events (PRO-AEs) in adjuvant and metastatic breast cancer: clinical and translational implications. Support Care Cancer 2017 Feb;25(2):549-558. [doi: 10.1007/s00520-016-3437-2] [Medline: 27747478]
23. Abernethy AP, Herndon JE, Wheeler JL, Day JM, Hood L, Patwardhan M, et al. Feasibility and acceptability to patients of a longitudinal system for evaluating cancer-related symptoms and quality of life: pilot study of an e/Tablet data-collection system in academic oncology. J Pain Symptom Manage 2009 Jun;37(6):1027-1038. [doi: 10.1016/j.jpainsymman.2008.07.011] [Medline: 19394793]
24. Pakhomov SV, Jacobsen SJ, Chute CG, Roger VL. Agreement between patient-reported symptoms and their documentation in the medical record. Am J Manag Care 2008 Aug;14(8):530-539 [FREE Full text] [Medline: 18690769]
25. Wuerstlein R, Kirkovits T, Drewes C, Schiltz D, Bauerfeind I, Haidinger R, et al. Abstract P2-10-05: eHealth in modern breast cancer treatment: new possibilities in communication between patients, doctors and nursing staff:. Cancer Res 2015 Apr 30;75(9 Supplement):Abstract nr P2-10-05. [doi: 10.1158/1538-7445.SABCS14-P2-10-05]
26. FDA (Food and Drug Administration). FDA. 2009. Guidance for industry: patient-reported outcome measures; use in medical product development to support labeling claims, Stand URL: https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM193282.pdf [accessed 2017-08-28] [WebCite Cache ID 6t3CzVnjD]

27. Snyder CF, Jensen RE, Segal JB, Wu AW. Patient-reported outcomes (PROs): putting the patient perspective in patient-centered outcomes research. Med Care 2013 Aug;51(8 Suppl 3):S73-S79 [FREE Full text] [doi: 10.1097/MLR.0b013e31829b1d84] [Medline: 23774513]
28. Bitton A, Onega T, Tosteson AN, Haas JS. Toward a better understanding of patient-reported outcomes in clinical practice. Am J Manag Care 2014 Apr;20(4):281-283 [FREE Full text] [Medline: 24884859]
29. Valderas JM, Alonso J. Patient reported outcome measures: a model-based classification system for research and clinical practice. Qual Life Res 2008 Nov;17(9):1125-1135. [doi: 10.1007/s11136-008-9396-4] [Medline: 18836850]
30. Szende A, Leidy NK, Revicki D. Health-related quality of life and other patient-reported outcomes in the European centralized drug regulatory process: a review of guidance documents and performed authorizations of medicinal products 1995 to 2003. Value Health 2005;8(5):534-548 [FREE Full text] [doi: 10.1111/j.1524-4733.2005.00051.x] [Medline: 16176492]
31. Basch E. Patient-reported outcomes in drug safety evaluation. Ann Oncol 2009 Dec;20(12):1905-1906 [FREE Full text] [doi: 10.1093/annonc/mdp542] [Medline: 19934250]
32. Basch E, Jia X, Heller G, Barz A, Sit L, Fruscione M, et al. Adverse symptom event reporting by patients vs clinicians: relationships with clinical outcomes. J Natl Cancer Inst 2009 Dec 2;101(23):1624-1632 [FREE Full text] [doi: 10.1093/jnci/djp386] [Medline: 19920223]
33. Basch E, Goldfarb S. Electronic patient-reported outcomes for collecting sensitive information from patients. J Support Oncol 2009;7(3):98-99. [Medline: 19507457]
34. Basch E. New frontiers in patient-reported outcomes: adverse event reporting, comparative effectiveness, and quality assessment. Annu Rev Med 2014;65:307-317.
35. Willke RJ, Burke LB, Erickson P. Measuring treatment impact: a review of patient-reported outcomes and other efficacy endpoints in approved product labels. Control Clin Trials 2004 Dec;25(6):535-552. [doi: 10.1016/j.cct.2004.09.003] [Medline: 15588741]
36. Velikova G, Booth L, Smith AB, Brown PM, Lynch P, Brown JM, et al. Measuring quality of life in routine oncology practice improves communication and patient well-being: a randomized controlled trial. J Clin Oncol 2004 Feb 15;22(4):714-724 [FREE Full text] [doi: 10.1200/JCO.2004.06.078] [Medline: 14966096]
37. Wallwiener M, Simoes E, Sokolov AN, Brucker SY, Fasching PA, Graf J. Health-related quality of life in metastatic and adjuvant breast cancer patients. Geburtshilfe Frauenheilkd 2016 Oct;76(10):1065-1073 [FREE Full text] [doi: 10.1055/s-0042-113188] [Medline: 27761027]
38. 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. [Medline: 8433390]
39. Basch E, Iasonos A, Barz A, Culkin A, Kris MG, Artz D, et al. Long-term toxicity monitoring via electronic patient-reported outcomes in patients receiving chemotherapy. J Clin Oncol 2007 Dec 01;25(34):5374-5380. [doi: 10.1200/JCO.2007.11.2243] [Medline: 18048818]
40. Judson TJ, Bennett AV, Rogak LJ, Sit L, Barz A, Kris MG, et al. Feasibility of long-term patient self-reporting of toxicities from home via the internet during routine chemotherapy. J Clin Oncol 2013 Jul 10;31(20):2580-2585 [FREE Full text] [doi: 10.1200/JCO.2012.47.6804] [Medline: 23733753]
41. Graf J, Simoes E, Wißlicen K, Rava L, Walter CB, Hartkopf A, et al. Willingness of patients with breast cancer in the adjuvant and metastatic setting to use electronic surveys (ePRO) depends on sociodemographic factors, health-related quality of life, disease status and computer skills. Geburtshilfe Frauenheilkd 2016 May;76(5):535-541 [FREE Full text] [doi: 10.1055/s-0042-105872] [Medline: 27239062]
42. Hartkopf AD, Graf J, Simoes E, Keilmann L, Sickenberger N, Gass P, et al. Electronic-based patient-reported outcomes: willingness, needs, and barriers in adjuvant and metastatic breast cancer patients. JMIR Cancer 2017 Aug 07;3(2):e11 [FREE Full text] [doi: 10.2196/cancer.6996] [Medline: 28784595]
43. Barentsz MW, Wessels H, van Diest PJ, Pijnappel RM, Haaring C, van der Pol CC, et al. Tablet, web-based, or paper questionnaires for measuring anxiety in patients suspected of breast cancer: patients' preferences and quality of collected data. J Med Internet Res 2014;16(10):e239 [FREE Full text] [doi: 10.2196/jmir.3578] [Medline: 25364951]
44. Steele GC, Gill A, Khan AI, Hans PK, Kuluski K, Cott C. The electronic patient reported outcome tool: testing usability and feasibility of a mobile app and portal to support care for patients with complex chronic disease and disability in primary care settings. JMIR Mhealth Uhealth 2016 Jun 02;4(2):e58 [FREE Full text] [doi: 10.2196/mhealth.5331] [Medline: 27256035]
45. Ayana BA, Negash S, Yusuf L, Tigeneh W, Haile D. Reliability and validity of amharic version of EORTC QLQ-C 30 questionnaire among gynecological cancer patients in Ethiopia. PLoS One 2016;11(6):e0157359 [FREE Full text] [doi: 10.1371/journal.pone.0157359] [Medline: 27304066]
46. Schwartzberg L. Electronic patient-reported outcomes: the time is ripe for integration into patient care and clinical research. Am Soc Clin Oncol Educ Book 2016;35:e89-e96 [FREE Full text] [doi: 10.14694/EDBK_158749] [Medline: 27249775]
47. Lachin JM. Statistical properties of randomization in clinical trials. Control Clin Trials 1988 Dec;9(4):289-311. [Medline: 3060315]

48. Schulz KF. Randomized controlled trials. Clin Obstet Gynecol 1998 Jun;41(2):245-256. [Medline: 9646957]
49. EORTC. 2001. EORTC QLQ C-30 scoring manual URL: http://www.eortc.be/qol/files/scmanualqlq-c30.pdf [accessed 2017-09-06] [WebCite Cache ID 6tGrOCdVA]
50. Brady MJ, Cella DF, Mo F, Bonomi AE, Tulsky DS, Lloyd SR, et al. Reliability and validity of the functional assessment of cancer therapy-breast quality-of-life instrument. J Clin Oncol 1997 Mar;15(3):974-986. [Medline: 9060536]
51. Kemmler G, Holzner B, Kopp M, Dünser M, Margreiter R, Greil R, et al. Comparison of two quality-of-life instruments for cancer patients: the functional assessment of cancer therapy-general and the European Organization for Research and Treatment of Cancer Quality of Life Questionnaire-C30. J Clin Oncol 1999 Sep;17(9):2932-2940. [doi: 10.1200/JCO.1999.17.9.2932] [Medline: 10561373]
52. Schaeffeler N, Pfeiffer K, Ringwald J, Brucker SY, Wallwiener M, Zipfel S, et al. Assessing the need for psychooncological support: screening instruments in combination with patients' subjective evaluation may define psychooncological pathways. Psychooncology 2015 Dec;24(12):1784-1791. [doi: 10.1002/pon.3855] [Medline: 26042392]

Abbreviations

CONSORT: Consolidated Standards of Reporting Trials
EORTC QLQ-C30: European Organization for Research and Treatment of Cancer Quality of Life Questionnaire-Core 30 item
e-PRO: electronic-based patient-reported outcome
e-PROCOM: Patient-Reported Outcomes and Compliance Analysis
FACT-G: Functional Assessment of Cancer Therapy-General
HRQoL: health-related quality of life
MD: median
PRO: patient-reported outcome
QoL: quality of life
SD: standard deviation
SPSS: Statistical Package for the Social Sciences version

Edited by G Eysenbach; submitted 28.06.17; peer-reviewed by A Birukov, N Schaeffeler; comments to author 13.07.17; revised version received 08.08.17; accepted 14.08.17; published 14.09.17.

Please cite as:
Wallwiener M, Matthies L, Simoes E, Keilmann L, Hartkopf AD, Sokolov AN, Walter CB, Sickenberger N, Wallwiener S, Feisst M, Gass P, Fasching PA, Lux MP, Wallwiener D, Taran FA, Rom J, Schneeweiss A, Graf J, Brucker SY
Reliability of an e-PRO Tool of EORTC QLQ-C30 for Measurement of Health-Related Quality of Life in Patients With Breast Cancer: Prospective Randomized Trial
J Med Internet Res 2017;19(9):e322
URL: http://www.jmir.org/2017/9/e322/
doi:10.2196/jmir.8210
PMID:28912116

Original Paper

Computerized Adaptive Testing Provides Reliable and Efficient Depression Measurement Using the CES-D Scale

Bao Sheng Loe[1*], Mphil; David Stillwell[2*], PhD; Chris Gibbons[2,3*], PhD

[1]School of Psychology, University of Cambridge, Cambridge, United Kingdom
[2]The Psychometrics Centre, Judge Business School, University of Cambridge, Cambridge, United Kingdom
[3]Cambridge Centre for Health Services Research, Cambridge School of Clinical Medicine, University of Cambridge, Cambridge, United Kingdom
[*]all authors contributed equally

Corresponding Author:
Chris Gibbons, PhD
The Psychometrics Centre
Judge Business School
University of Cambridge
Trumpington Street
Cambridge,
United Kingdom
Phone: 44 01223765203
Email: cg598@cam.ac.uk

Abstract

Background: The Center for Epidemiologic Studies Depression Scale (CES-D) is a measure of depressive symptomatology which is widely used internationally. Though previous attempts were made to shorten the CES-D scale, few have attempted to develop a Computerized Adaptive Test (CAT) version for the CES-D.

Objective: The aim of this study was to provide evidence on the efficiency and accuracy of the CES-D when administered using CAT using an American sample group.

Methods: We obtained a sample of 2060 responses to the CESD-D from US participants using the myPersonality application. The average age of participants was 26 years (range 19-77). We randomly split the sample into two groups to evaluate and validate the psychometric models. We used evaluation group data (n=1018) to assess dimensionality with both confirmatory factor and Mokken analysis. We conducted further psychometric assessments using item response theory (IRT), including assessments of item and scale fit to Samejima's graded response model (GRM), local dependency and differential item functioning. We subsequently conducted two CAT simulations to evaluate the CES-D CAT using the validation group (n=1042).

Results: Initial CFA results indicated a poor fit to the model and Mokken analysis revealed 3 items which did not conform to the same dimension as the rest of the items. We removed the 3 items and fit the remaining 17 items to GRM. We found no evidence of differential item functioning (DIF) between age and gender groups. Estimates of the level of CES-D trait score provided by the simulated CAT algorithm and the original CES-D trait score derived from original scale were correlated highly. The second CAT simulation conducted using real participant data demonstrated higher precision at the higher levels of depression spectrum.

Conclusions: Depression assessments using the CES-D CAT can be more accurate and efficient than those made using the fixed-length assessment.

(J Med Internet Res 2017;19(9):e302) doi:10.2196/jmir.7453

KEYWORDS

depression; assessment; psychometrics; patient reported outcome measures; patient outcome assessment

Introduction

The Center for Epidemiologic Studies Depression Scale (CES-D) is a commonly used 20-item self-rating scale designed to measure depressive symptomatology in both clinical and non-clinical settings [1]. It is used in both epidemiological research and as a diagnostic screening tool [2,3].

Despite much debate on the cut-off score which yield better sensitivities and specificities [3,4], it is commonly accepted that persons who score 16 or above on the CES-D's 0 to 60 scale are likely to be clinically depressed [5,6]. While some authors have suggested that the CES-D has a four-factor structure, it appears to provide meaningful measurement along a single dimension [7]. Hence, this level of internal consistency suggests that the CES-D should be used as an overall scale to measure a single latent construct—depressive symptoms [8].

Although the fixed-length version of the CES-D is widely used, recent developments in the availability of software to conduct advanced psychometric analyses and to develop computer adaptive assessments bring new opportunities for advanced Internet-based depressive symptom assessment. Computerized Adaptive Testing (CAT), refers to an algorithm-based assessment protocol which iteratively matches participants in a psychometric assessment with the most relevant item for them. Conducting assessments in this manner often reduces the number of items which need to be administered in an assessment, reducing the length of assessments by as much as 82%, compared to fixed-length measures of the same construct [9-11]. CAT typically relies on item parameter information derived from item-response theory. A large number of item-response theory models are suitable for developing item banks including the graded response model (GRM), the Rasch family of models as well as multidimensional models [12,13].

As well as demonstrable increases in efficiency, CATs can deal with other issues which prohibit accurate measurement using static questionnaires. For example, CATs are able to adjust for demographic differences in the interpretations of items commonly seen between different groups and known as differential item functioning (DIF) [14-16]. It is also possible to account for issues caused by items being *too* similar which can spuriously inflate assessment reliability [17].

An investigation conducted by Smit et al [10] demonstrated that the CES-D items make suitable candidates for CAT administration in a sample of Dutch adolescents aged between 12 and 17 [18]. The study shows that CAT administration could approach the reliability of the paper-based measures using fewer than half the items on the original. Other CATs have developed novel item banks to create CATs of depression, including the D-CAT [19,20] and PROMIS depression item banks [21]. These item banks show similar performance, arriving at reliable estimates of depressive symptomology using fewer than 10 items. Though both using legacy questionnaires to “feed” CATs and developing item banks specifically for that purpose have advantages and disadvantages. One advantages of using the CES-D for CAT is that it is not only well known and widely understood but it is also freely available in the public domain, allowing its use as a CAT assessment without incurring additional fees or reliance on restrictive proprietary software.

Thus, this paper aims to validate the CES-D assessment for use as a Web-based CAT using a sample taken from the US general population which will allow patients, clinicians, and other members of the public to evaluate depression symptomology efficiently and precisely online.

Methods

Participants

We recruited 2060 individuals who completed the CES-D scale via the myPersonality application [22]. MyPersonality is a Facebook application that allowed Facebook users to complete psychological tests and receive feedback on their scores. Users of the myPersonality application provided opt-in consent to allow us to record their assessment scores in exchange for the opportunity to receive feedback, which can be later shared online. The sample was divided into two groups using a randomly generated numeric string (random.org) for analysis. The first group is used for evaluation of the CES-D scale (n=1018). The second group is used for validating the CAT results based the calibration of the item parameters derived from the evaluation sample (n=1042). The samples were independent from one another. For group 1, there were 65.52% (665/1018) 6 females and 34.39% (348/1012) males. The mean age of the participants was 26 years (SD 12.12). For group 2, there were 65.93% (687/1042) females and 33.69% (351/1042) males. The mean age for participants was 25.86 (SD 10.44). Five participants from group 1 and 4 participants from group 2 did not reveal their gender. All individuals reported that they were from the United States.

Measure

The CES-D is a self-report questionnaire which measures severity of depression from the perspective of the individual (see Multimedia Appendix 1). Subjects responded to the CES-D by indicating on a 4-point Likert-scale stating how often each depressive symptom occurred during the past week (0=rarely or none of the time, 1=some of the time, 2=much of the time, 3=most or all the time). The potential range of scores is from 0 to 60, with higher scores indicating higher levels of depressive symptomology.

The CES-D scale is a well validated and widely used instrument in many studies internationally [23-25]. Reliability and validity of the scale has been tested in both general and clinical populations [1]. Previous results show that the 20-item scale yields good internal consistency for the general population (Cronbach alpha=.85) and for a psychiatric population (Cronbach alpha=.90) [26]. Adequate test-retest reliability was found over 2 to 8-week period and 3 to 12-month period, respectively [26,27]. Convergent validity was supported by the significant correlations with other scales designed to assess depression symptoms [18,19,28,29]. The CES-D scale is available to use in the public domain and free to use without restriction.

Data Analysis

The internal consistency of the CES-D scale was determined using the Cronbach alpha statistic [30], confirmatory factor analysis (CFA) was first performed to determine the structure of the model. The maximum-likelihood estimator was in the confirmatory analyses. Four fit indices were used in this study: chi-square statistics [31]; the Comparative fit index (CFI, [32]), the Tucker Lewis Index (TLI, Tucker and Lewis 1973), and the root-mean-square error of approximation (RMSEA, [33]). The chi-square statistics indicates whether the observed covariance matrix is similar to the predicted covariance matrix. However, the result is liable to bias in large sample sizes [34]. As such, other criteria such as absolute and comparative fix indices are used to evaluate the model. The CFI and TLI indices are the relative reduction in lack of fit of an observed model versus an independent model; with values of 0.90 or greater indicating an adequate fit [35]. For RMSEA [33], values less than 0.05 indicate good fit, and values greater than 0.10 as indication of poor fit of a model after accounting for degrees of freedom of the model.

Subsequently, Mokken analysis was used to provide further insight into the scale's factor structure and the scalability of the items [36,37]. Following Mokken analysis, data were analyzed using GRM [38], which has been shown to be suitable for calibrating items for use as CAT assessments [39]. Item discrimination values ranging from 0.64 to 1.34 were considered to be moderately discriminative, and values 1.35 or greater are highly discriminative [40].

Following the protocol set out by the PROMIS investigators [41], we assessed the assumptions of GRM and made modifications, where necessary, to the scale to resolve breaches of model assumptions, which are detailed below.

Local independence of items was assessed using Yen's Q3 method of correlated residuals. Item residual correlations above .20 were considered indicative of local dependence between items [42]. Different strategies exist for managing items with local dependency, which including removing the items from the scale completely or collapsing the items into a testlet.

The DIF analysis using the lordif package was conducted for age and gender groups to identify measurement biases between groups [43]. The lordif package utilizes ordinal logistic regression methods to calculate DIF [44]. DIF is observed when the probability of answering a specific item correctly is not the same for individuals with the same level of depressive symptoms but who belong to a different demographic group [15]. For example, male and female participants may both have equal levels of depressive symptoms, but if the certain items are interpreted differently between groups then observed mean scores may incorrectly show that one group has higher levels of depressive symptoms than the other because of an artefact of their gender that was not adequately controlled for within the test. Hence, DIF is used to identify items with unwanted bias and indicate that the same item sets and parameters might be needed for different diagnostic groups [16].

We conducted DIF analysis to assess item invariance with respect to age and gender. Two criteria were adopted in this study to detect meaningful DIF: changes in the beta [43] and the pseudo R-square [45]. Values ranging from 5% to 10% beta change and pseudo R-squared >.13 suggest that meaningful DIF exist for a particular item [43,45,46]. For our study, items with beta change of above >1% was flagged for DIF. We divided the sample into 2 groups based on the mean age (26 years) of the sample. Participants who were younger than mean age were placed in the first group (n=399) and those that were older than the mean age were placed in the second group (n=200). For gender groups, all the males were in the first group (n=348), whereas all the females were in second group (n=665). Participants who did not wish to reveal their gender (n=9) were excluded from the DIF analysis as there were too few to create an adequate additional group.

We evaluated the impact of DIF on the CES-D scores by recalibrating the items to the GRM model using the DIF-adjusted item parameters [47]. The person scores were recalculated based on these parameters. Finally, the strength of the association between the DIF-adjusted person score and original person score were evaluated using Pearson correlation. A high correlation would suggest that adjusting for DIF would make negligible differences in the person scores, and as such, could be ignored [48]. A low correlation between the DIF-adjusted person scores and original person scores suggest that the DIF makes a meaningful difference on the final scores and that group-specific parameters should be used when developing a CAT.

Establishing Evaluation of CES-D CAT Simulation

Two simulations were conducted to evaluate the properties of the item pool and the CAT algorithm. The first simulation employed simulated responses from various levels of the latent trait derived from participants who completed the full CES-D scale to determine the average number of items that had to be administered.

The second simulations were respondents from the validation group and thus, the simulations were conducted using real data. The item parameter estimates used in the CES-D CAT were derived from the evaluation group. The validation sample used in this simulation did not overlap with the evaluation sample used to calibrate the item bank. Nevertheless, the individuals of this sample completed the same CES-D items that had been employed in the construction of the item bank. As such, responses to all items in the item bank were available. Both the respondents' latent trait levels and responses to individual items were used to estimate the number of items needed to administer in a CAT. Correlations with the simulated CAT score and their scores derived from the full CES-D were obtained for both groups.

The maximum Fisher Information criterion was used for item selection [49,50]. The Bayesian modal estimation was used at the beginning of the CAT simulation to estimate ability [51]. This approach temporarily assumes that the ability of the test takers is normally distributed. Once a mixed response pattern is obtained, the normal distribution assumption is no longer requires and thus, a non-Bayesian maximum likelihood estimation is used [52]. Maximum likelihood estimation is subsequently used to estimate the final ability of the test taker [49]. The major advantage of using maximum likelihood estimation of ability is that it can account for all the information in the test taker's responses in conjunction with the information available on each test item. The stopping rule for both simulations were set at SE≤0.32, which roughly corresponds to a reliability value≥0.90 [53].

Software

Analyses were all conducted using the R Statistical Computing Environment [54]. Individual packages were loaded to conduct CFA ("lavaan," [55]), Mokken ("mokken," [56]) and item response theory (IRT) including CAT simulations (“mirt,” [57] and “catR” [Magis and Raîche, 2011]).

Results

Confirmatory Factor Analysis

Confirmatory factor analysis (CFA) was employed to investigate the unidimensionality of the CES-D scale. Table 1 lists the mean, standard deviation, and the factor loadings of the CES-D items, revealing no reason for concern about the multivariate distribution of the data. Therefore, the model was estimated using the maximum likelihood method. As shown, the factor loadings are above the recommended threshold of .3 (Kline, 2013).

Initial CFA results indicate a poor fit to the model ($\chi^2_{8.4}$, $P<.05$; TLI=0.94; CFI=0.86; and RMSEA=0.09 (95% CI=0.08-0.09)).

Unidimensionality

We used Mokken analysis to further explore the dimensional structure of the CES-D and identify the potential sources of multidimensionality identified with the CFA. The evaluation of item homogeneity is based on the Loevinger’s *H* coefficient [58]. Scalability is considered to be sufficient for both items and the scale where Loevinger’s *H* is equal to or greater than 0.30 [59]. We found that items 2, 11, and 15 displayed item coefficients of homogeneity<0.3. Hence, these items were eliminated from further analysis. This strategy was repeated and all the items were found to be above the recommended threshold, which conformed to a single dimension with Loevinger’s coefficient of homogeneity at a scale level of 0.43 (Table 2).

Table 1. Factor loadings and item descriptive statistics for the CES-D scale.

Item no.	Mean	SD	Factor loadings
q1	2.09	0.95	0.53
q2	1.94	1.04	0.40
q3	2.22	1.07	0.81
q4	2.34	1.06	0.58
q5	2.67	1.00	0.51
q6	2.42	1.04	0.85
q7	2.48	1.01	0.49
q8	2.43	0.99	0.56
q9	2.16	1.09	0.70
q10	2.09	1.03	0.54
q11	2.57	1.11	0.42
q12	2.31	0.94	0.72
q13	2.25	1.01	0.60
q14	2.76	1.07	0.69
q15	1.90	0.92	0.41
q16	2.36	0.98	0.71
q17	1.77	0.96	0.54
q18	2.59	0.99	0.80
q19	2.30	1.07	0.63
q20	2.51	1.01	0.58

Table 2. Loevinger's coefficient of homogeneity at an item-level.

Item	Mean	Item H (H_i)[a]	Standard Error	Dimensionality
1	2.09	0.38	0.02	1
3	2.22	0.55	0.01	1
4	2.34	0.41	0.02	1
5	2.67	0.38	0.02	1
6	2.42	0.57	0.01	1
7	2.48	0.35	0.02	1
8	2.43	0.39	0.02	1
9	2.16	0.49	0.02	1
10	2.10	0.39	0.02	1
12	2.31	0.50	0.02	1
13	2.25	0.42	0.02	1
14	2.76	0.48	0.02	1
16	1.90	0.49	0.02	1
17	2.36	0.40	0.02	1
18	1.77	0.55	0.01	1
19	2.60	0.43	0.02	1
20	2.30	0.41	0.02	1

[a] Scale H=0.45.

Graded Response Model

Once we have established unidimensionality using Mokken analysis. We fitted the remaining 17 items to Samejima's GRM (Table 3). The slope and threshold parameters in the GRM are used describe the relationship between each item and overall depressive symptom severity. The slope parameter reflects how well the items discriminate between respondents with or without depressive symptoms. The item discrimination values (alpha) ranged from a high of alpha=3.70 (item 5) to a relative low, but still strong, alpha=1.13 (item 7). The threshold parameter describes the endorsement of depressive symptoms, with larger values indicating greater levels of depressive symptoms. The thresholds for the lowest item category (b1) ranged from −3.31 (item 18) to 0.15 (item 17) on a z-score scale, indicating low to average levels of depressive symptoms, relative to the rest of our sample, for the individuals who endorsed the lowest CES-D category. The thresholds for the highest CES-D category (b3) ranged from 3.50 (item 3) to 1.25 (item 14), indicating moderate to high levels of depressive symptoms. All the standard errors of the *b* estimates were considered marginal, indicating that the items were normally distributed. An item fit analysis was conducted to identify any misfits. However, the results indicated that the remaining items fitted the model. Examination of the factor loadings revealed that all items loaded significantly (>.50) on the single factor. Therefore, this model described the data adequately.

Local independence

Local dependency was apparently between items 18,12, and 16 as well as items 1, 19, and 15. Items 8, 12, and 16 were grouped as testlet 1, and items 19, and 15 were grouped as testlet 2. We observed the item residual correlation and found that item 4 was still correlated (>0.2) with the first testlet. Hence, we grouped item 4 together with the first testlet and repeated the analysis, resulting in no correlated residuals greater than 0.2. Within the IRT framework, the fit indices based on the limited information M_2 statistic was used to assess the model fit [60]. The result shows that the RMSEA was at 0.065 (95% CI 0.06-0.07), and comparative indices (TLI=0.96, CFI=0.97) were above the recommended threshold [35].

Figure 1 displays the test information curve for the IRT GRM. The test characteristics curve is simply the additive of the scores associated with increasing levels of depressive symptoms. The test information is at its highest (18.71) when the theta level is slightly above 0, while the lowest around of information can be found at both tails of the *x*-axis. Hence, the CES-D scale is most precise in estimating the underlying trait when the theta level is approximately zero (average).

Table 3. Parameter estimates and factor loadings for the 17 items of the CES-D Scale.

Item	a	b1	b2	b3	Factor 1
Item 1	1.22 (0.09)	−0.97 (0.09)	0.93 (0.09)	2.85 (0.15)	0.58
Item 3	3.15 (0.18)	−1.56 (0.15)	1.14 (0.15)	3.50 (0.21)	0.88
Item 4	1.45 (0.09)	−1.31 (0.10)	0.20 (0.09)	2.16 (0.12)	0.65
Item 5	1.14 (0.08)	−2.13 (0.15)	−0.38 (0.08)	1.44 (0.10)	0.56
Item 6	3.70 (0.22)	−2.81 (0.21)	0.37 (0.16)	3.70 (0.24)	0.91
Item 7	1.13 (0.07)	−1.71 (0.10)	0.04 (0.08)	1.81 (0.10)	0.55
Item 8	1.37 (0.09)	−1.84 (0.11)	0.19 (0.09)	2.11 (0.12)	0.63
Item 9	2.03 (0.12)	−0.81 (0.11)	0.86 (0.11)	2.66 (0.15)	0.77
Item 10	1.29 (0.09)	−0.66 (0.09)	0.79 (0.09)	2.56 (0.13)	0.60
Item 12	2.14 (0.13)	−2.09 (0.15)	0.65 (0.12)	3.33 (0.18)	0.78
Item 13	1.49 (0.09)	−1.27 (0.10)	0.58 (0.09)	2.49 (0.13)	0.66
item 14	1.98 (0.11)	−2.59 (0.14)	−0.70 (0.11)	1.25 (0.11)	0.76
Item 16	2.12 (0.13)	−2.06 (0.14)	0.44 (0.12)	2.98 (0.16)	0.78
Item 17	1.38 (0.10)	0.15 (0.09)	1.57 (0.11)	3.24 (0.18)	0.63
Item 18	2.89 (0.16)	−3.31 (0.20)	−0.34 (0.14)	2.82 (0.18)	0.86
Item 19	1.50 (0.10)	−1.21 (0.10)	0.41 (0.09)	2.15 (0.12)	0.66
Item 20	1.37 (0.09)	−1.99 (0.12)	0.09 (0.09)	1.83 (0.11)	0.63

Figure 1. The test information of the CES-D scale.

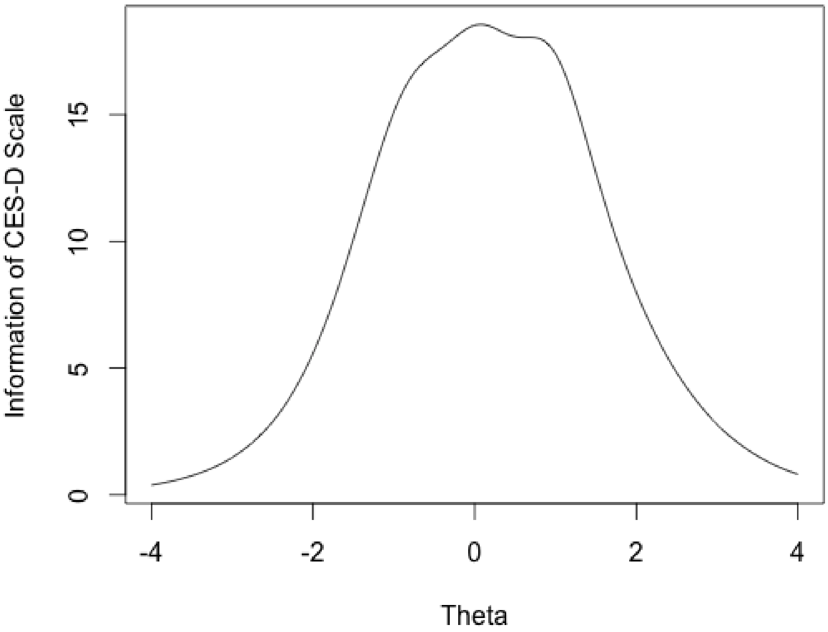


Table 4. CES-D CAT Simulation of respondents.

Measure	D1[a]	D2	D3	D4	D5	D6	D7	D8	D9	D10
Mean theta	−1.70	−1.04	−0.69	−0.37	−0.09	0.16	0.40	0.68	1.01	1.64
RMSE[b]	0.29	0.35	0.31	0.33	0.32	0.32	0.30	0.32	0.34	0.34
Mean bias	0.01	0.00	0.02	0.00	0.02	0.02	−0.03	−0.05	−0.07	0.01
Mean test length	13.98	12.34	11.04	9.50	8.30	7.70	7.60	8.18	9.77	13.09
Mean standard error	0.34	0.34	0.33	0.34	0.33	0.33	0.33	0.34	0.34	0.34
Number of simulees	105	104	104	104	104	104	104	104	104	105

[a] D: decile.

[b] RMSE: root mean square error

DIF Analysis

DIF was not found between age groups. However, results indicated that item 14 (“I felt lonely”) showed moderate DIF for gender groups, with a beta change of more than 1% and a pseudo R-square of 0.08. When the DIF-adjusted person scores were calculated, the Pearson correlation between the original person scores and the DIF-adjusted person scores were 0.99. A *t* test analysis showed a non-significant mean difference with scores for DIF-adjusted person scores (mean=0.01, SD=0.88), and original person scores (mean=0.00, SD=0.97); $t_{2016.4}$=−0.15, *P*=.88.

On the basis of these results, the conclusion arrived at was that statistically significant DIF was identified for item 14 using the two criteria of beta change and pseudo R-squared. However, the strength of association between the original person scores and the DIF-adjusted person scores were greater than 0.99. Therefore, the final decision was that any DIF found between the groups could be disregarded.

Computer Adaptive Testing Simulation

Simulation I

Table 4 presents the results of the first simulation. In this analysis, the data were sorted into 10 equal parts, with each part representing one tenth of the sample group. There are appropriate 104 or 105 participants in each decile (D) rank (n=1042). The estimated average test length was 10.16 with SD of 2.34. The mean RMSE was .32 and the mean bias was −.0083. The lowest number of items administered to the simulees was in D5, with an average of 8.3 test items. The lowest and highest decile rank requires substantially more items (D1=13.98 items; D10=13.09 items) in order to reach the same target precision of SE≤0.32.

Simulation II

The second simulation study was conducted using the responses from a sample of real respondents (validation group) who completed the full CES-D scale. The stopping rule was set at SE≤0.32. The result of the second simulation can be found in Table 5. On average, 11.72 items with SD of 2.68 were required to estimate the latent trait at this level of precision. The mean RMSE was 1.14 and the mean bias was 0.18. Unlike the first simulation, only respondents in the lowest decile ranking required the administration of substantially more items to reach the specified level of precision (D1=14.76 items). Interestingly, there is a downward trend in the length of items from D9 (14.61 items) to D10 (8.41 items), indicating higher precision at higher levels of depressive symptoms with the use of lesser items.

Further inspection of the item administration pattern (Figure 2), suggests a drop in the number of items required to estimate the latent trait accurately around. This could be due of the CAT algorithm selecting items with the highest information at every step, resulting in a quicker estimate of the latent trait. Figure 2 shows the number of items administered by the CES-D CAT as a function of the standardized score of the depressive symptoms construct.

Estimates of the level of CES-D trait score provided by the simulated CAT algorithm and the original CES-D trait score derived from original scale correlated highly (*r*=0.98). This indicates that a precise estimation of the latent trait is possible with substantial item savings using CAT approaches (Figure 3).

Figure 3 shows exceptionally high correlation between the score from the CAT and the score given to the same participants when every item was completed.

Table 5. CES-D CAT Simulation of CAT algorithm.

Measure	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10
Mean Theta	−1.70	−1.04	−0.69	−0.37	−0.09	0.16	0.40	0.68	1.01	1.64
RMSE	1.46	1.15	0.86	0.42	0.36	0.50	0.61	0.92	1.65	2.08
Mean bias	−1.45	−1.13	−0.82	−0.29	0.11	0.34	0.53	0.85	1.58	2.05
Mean test length	14.76	14.04	13.44	12.81	9.67	8.17	8.70	12.55	14.61	8.41
Mean standard error	0.42	0.34	0.33	0.33	0.33	0.33	0.34	0.34	0.33	0.37
Number of simulees	105	104	104	104	104	104	104	104	104	105

Figure 2. Relationship between number of items administered and level of depression (theta).

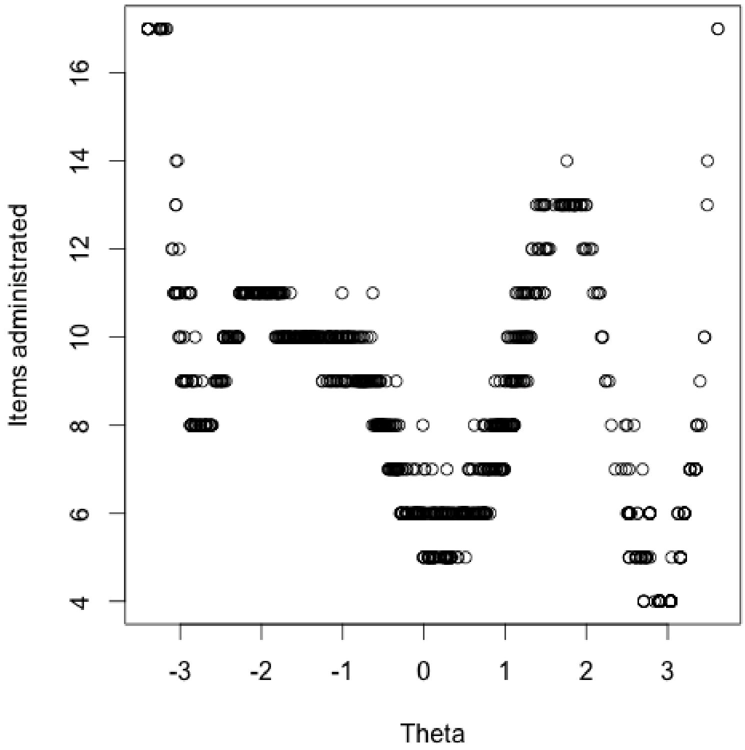
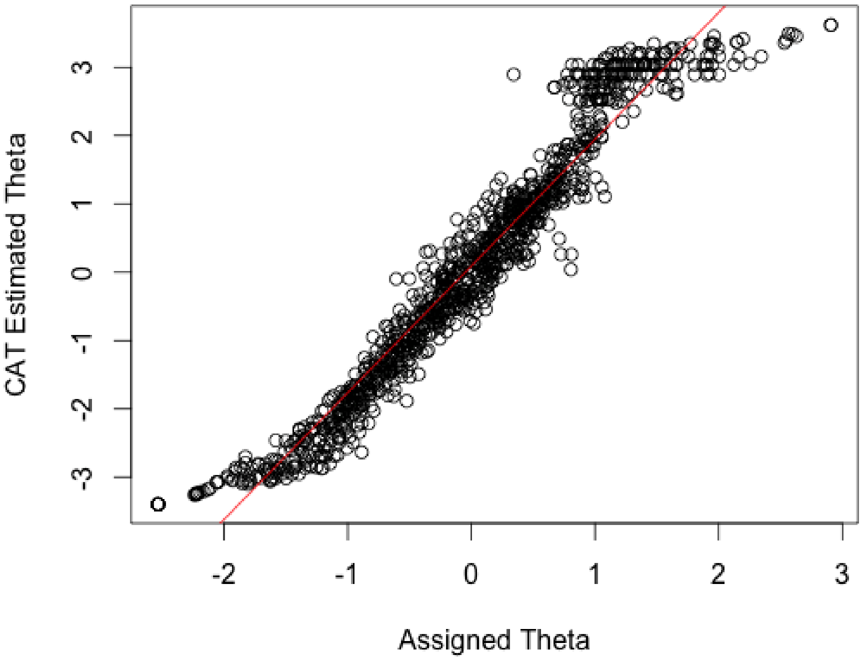


Figure 3. Comparison of CES-D CAT scores with an IRT score computed from all items in the item bank.



Discussion

Principal Findings

The psychometric properties of the CES-D measure were evaluated using a US sample. This sample was chosen to avoid issues of DIF across culture, and with the aim of providing an item bank which could be suitable for use within a clinical and research setting in the United States. The CES-D scale displayed excellent internal consistency based on the Cronbach alpha. The factor structure of the CES-D was subsequently evaluated using confirmatory factor and Mokken analysis. However, the results from the CFA indicated that the model provided an inadequate fit to the data. Mokken analysis identified three items as sources of multidimensionality in the CES-D. Items 2, 11, and 15 were considered to have poor fit and were subsequently removed from the analysis. Item 2 referred to "I did not feel like eating; my appetite was poor," item 11 referred to "my sleep was restless," and item 15 referred to "People were unfriendly." The result showed that the final 17 items were found to be suitable for measuring a unidimensional trait and thus, the item parameters achieved from the remaining items allowed us to develop a computerized adaptive CES-D assessment.

The CES-D scale was calibrated using the IRT approaches. Most IRT based models require that items measure a single underlying dimension and this condition was met based on the result of the Mokken analysis. Furthermore, IRT based frameworks made computer-adaptive CES-D possible with the estimated item parameters derived from IRT models. Simulated computer-adaptive administration of the item bank demonstrates the ability to estimate precise latent trait levels with similar or higher levels of internal reliability similar to the original scale but using fewer items. These results are commensurate with other research exploring the performance of the CES-D as a CAT in other contexts including adolescents and people with multimorbidity [10,61] and for adaptive testing of depressive symptoms using the PROMIS system [62].

Unlike a test developed using classical test theory, in which the number of items is fixed and precision naturally varies between participants who have differing levels of latent ability, CAT fixes the precision while allowing the number of items to vary. CAT can only be conducted using computer administration and the items are previously calibrated with a suitable item response model. The steps to conducting a CAT are (1) administer an item, (2) compute the latent score and its standard error, (3) identify the next most informative item based on the current latent score estimate and IRT parameters, and (4) repeat steps 1-3 until the predefined stopping rule has been met.

In our simulation studies, we found a very high correlation between the CAT scores obtained when all 17 items were administered and when the stopping rule was introduced (leading to a mean test length of 10 items). Moreover, at the extreme (higher) end of the latent trait continuum, it only requires 8 items to identify individuals with depressive symptoms. This encourages quicker assessment of depressive symptoms, which can help clinicians to identify potential groups of persons who may benefit from immediate medical intervention. In spite of substantial improvements in efficiency by employing the CAT procedure, little information is lost and scores are still estimated accurately. By comparison, the time taken to complete the CES-D CAT will be shorter than the original 20-item scale. This time saving may seem small as far as a single scale is concerned, but psychometric assessment usually involves multiple questionnaires and, from this perspective, substantial time saving is evident.

Limitations and Future Research

A limitation to the current research is the small number of items used to measure CES-D. With CAT, the precision of latent trait estimates increases with the number of items in the item bank. A smaller item bank gives fewer options for item selection and may results in reduced item variation between assessments. However, to apply stricter stopping-rule criteria means that the number of items necessary to complete the CES-D CAT will be about the same as completing all the original scale, thus, no extra benefit remains with the use of CES-D CAT. Therefore, while this study reports the stopping rule at less than or equal to 0.32, which is equivalent to a reliability of more than or equal to 0.90, the precision can still be heightened by increasing the test information. This can be achieved by adding more high-quality items to the item bank. Hence, future studies could evaluate the CAT system where new items are included as part of the test to increase the item bank and ensure that the performance of the CAT system is not compromised. The performance of the CAT algorithms can also be evaluated under "live" testing conditions rather than simulation of existing data to ensure that participants' test performance under conventional 'fixed length' and adaptive conditions do not differ significantly.

Compared with population-based samples used in the development of item banks elsewhere [21], our sample was younger and had a greater proportion of women. Given the nature of the recruitment into the study via a voluntary online app it is not surprising that this sample is more reflective of a "digitally native" population of younger people. One important caveat of this research is therefore that our findings should not be extrapolated to a general population but rather support the growing body of literature demonstrating the suitability of the CES-D for adaptive testing in different group as a means of making measurement more precise and efficient while retaining an item bank that is familiar to clinicians.

In this study, we assess the content validity of the CAT-administered CESD by comparing depressive symptom estimates from the full-length assessment with an adaptively administered version. Further research is required to establish to predictive validity of this tool for the correct classification of clinical depression to support its use in clinical contexts.

There are some discussions about the factor structure of the CES-D and whether a single factor is appropriate for assessing depressive symptomatology. Several researchers have suggested that the CES-D scale is a measure of the underlying 4-factor structure [1,63,64]. However, the construction of a 4-factor scale may be too challenging as psychometric test designed for health assessment aims to be as short as possible. Nevertheless, in the event that a 4-factor scale is developed for the CES-D, then a 4-factor CES-D CAT under the conditions of content balancing may be introduced. In other words, a proportionate sampling of items is taken from each of the factor domains, while ensuring unidimensionality is achieved [52]. Researchers can thus consider new research avenues in which one could understand in finer gradient of the depressive symptoms.

Conclusions

Our findings presented in this study shows that the CES-D CAT is a precise and efficient tool for screening depressive symptoms. Furthermore, the measurements provided by CAT are more likely to result in more meaningful research conclusions than classical approaches. More informed decisions could also be made based on measurement data at an individual level rather than at a scale level.

While increased complexity with regards to the test development is inevitable, the CES-D CAT has immediate advantages such as increased accuracy, exact interpretable, and shorter time spent over conventional testing approaches. Open source software such as the *Concerto* testing platform [65] makes it more accessible than ever before for researchers to develop and implement their own CAT system. Furthermore, the CESD-CAT outperforms the paper-based versions of the CES-D in terms of reliability, length, and flexibility in which they may be administered in a clinical setting. CATs are more dynamic as they adjust accordingly to the ability level of the test taker, indicating both efficiency and effectiveness. Thus, the CES-D CAT is suitable to be administered as a primary tool for understanding and screening individuals in the US with depressive symptomatology.

Conflicts of Interest

None declared.

Multimedia Appendix 1

CES-D questions.

[PDF File (Adobe PDF File), 12KB - jmir_v19i9e302_app1.pdf]

References

1. Radloff L. The CES-D scale a self-report depression scale for research in the general population. Appl Psychol Meas 1977;1(3):385-401. [doi: 10.1177/014662167700100306]
2. Himmelfarb S, Murrell SA. Reliability and validity of five mental health scales in older persons. J Gerontol 1983;38(3):333-339. [Medline: 6841929]
3. Myers JK, Weissman MM. Use of a self-report symptom scale to detect depression in a community sample. The American Journal of Psychiatry 1980;137(9):1081-1084. [Medline: 7425160]
4. Roberts RE. Reliability of the CES-D scale in different ethnic contexts. Psychiatry Res 1980;2(2):125-134. [Medline: 6932058]
5. Comstock GW, Helsing KJ. Symptoms of depression in two communities. Psychol Med 1977;6(4):551-563. [Medline: 1005571]
6. Hankin JR, Locke BZ. The persistence of depressive symptomatology among prepaid group practice enrollees: an exploratory study. Am J Public Health 1982;72(9):1000-1007. [Medline: 7102848]
7. Shafer AB. Meta - analysis of the factor structures of four depression questionnaires: Beck, CES - D, Hamilton, and Zung. J Clin Psychol 2006;62(1):123-146 [FREE Full text] [Medline: 16287149]
8. Kohout FJ, Berkman LF, Evans DA, Cornoni-Huntley J. Two shorter forms of the CES-D (Center for Epidemiological Studies Depression) depression symptoms index. J Aging Health 1993 May;5(2):179-193. [doi: 10.1177/089826439300500202] [Medline: 10125443]
9. Gibbons C, Bower P, Lovell K, Valderas J, Skevington S. Electronic quality of life assessment using computer-adaptive testing. J Med Internet Res 2016 Sep 30;18(9):e240 [FREE Full text] [doi: 10.2196/jmir.6053] [Medline: 27694100]
10. Smits N, Cuijpers P, van Straten A. Applying computerized adaptive testing to the CES-D scale: a simulation study. Psychiatry Res 2011 Jun 30;188(1):147-155 [FREE Full text] [doi: 10.1016/j.psychres.2010.12.001] [Medline: 21208660]
11. Hays RD, Morales LS, Reise SP. Item response theory and health outcomes measurement in the 21st century. Medical care 2000;38(9 Suppl):1128-1142 [FREE Full text] [Medline: 10982088]
12. Gibbons RD, Weiss DJ, Frank E, Kupfer D. Computerized adaptive diagnosis and testing of mental health disorders. Annu Rev Clin Psychol 2016;12:83-104. [doi: 10.1146/annurev-clinpsy-021815-093634] [Medline: 26651865]
13. Wahl I, Löwe B, Bjorner JB, Fischer F, Langs G, Voderholzer U, et al. Standardization of depression measurement: a common metric was developed for 11 self-report depression measures. J Clin Epidemiol 2014;67(1):86. [Medline: 24262771]
14. Bee P, Gibbons C, Callaghan P, Fraser C, Lovell K. Evaluating and quantifying user and carer involvement in mental health care planning (EQUIP): co-development of a new patient-reported outcome measure. PLoS One 2016;11(3):e0149973 [FREE Full text] [doi: 10.1371/journal.pone.0149973] [Medline: 26963252]
15. Holland PW, Wainer H. Differential item functioning. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc; 2012.
16. Tennant A, Penta M, Tesio L, Grimby G, Thonnard J, Slade A, et al. Assessing and adjusting for cross-cultural validity of impairment and activity limitation scales through differential item functioning within the framework of the Rasch model: the PRO-ESOR project. Med Care 2004 Jan;42(1 Suppl):I37-I48. [doi: 10.1097/01.mlr.0000103529.63132.77] [Medline: 14707754]
17. Wright BD. Local dependency, correlations and principal components. Rasch Measurement Transactions 1996;10(3):509-511.
18. Smits N, Cuijpers P, van Straten A. Applying computerized adaptive testing to the CES-D scale: a simulation study. Psychiatry Res Internet 2011;188(1):147-155 [FREE Full text] [Medline: 21208660]

19. Fliege H, Becker J, Walter OB, Rose M, Bjorner JB, Klapp BF. Evaluation of a computer-adaptive test for the assessment of depression (D-CAT) in clinical application. Int J Methods Psychiatr Res 2009;18(1):23-36. [doi: 10.1002/mpr.274] [Medline: 19194856]
20. Fliege H, Becker J, Walter OB, Bjorner JB, Klapp BF, Rose M. Development of a computer-adaptive test for depression (D-CAT). Qual Life Res 2005 Dec;14(10):2277-2291. [doi: 10.1007/s11136-005-6651-9] [Medline: 16328907]
21. Pilkonis PA, Choi SW, Reise SP, Stover AM, Riley WT, Cella D, PROMIS Cooperative Group. Item banks for measuring emotional distress from the patient-reported outcomes measurement information system (PROMIS®): depression, anxiety, and anger. Assessment 2011 Sep;18(3):263-283 [FREE Full text] [doi: 10.1177/1073191111411667] [Medline: 21697139]
22. Kosinski M, Matz SC, Gosling SD, Popov V, Stillwell D. Facebook as a research tool for the social sciences: opportunities, challenges, ethical considerations, and practical guidelines. Am Psychol 2015;70(6):543-556. [doi: 10.1037/a0039210] [Medline: 26348336]
23. Quiñones AR, Thielke SM, Clark ME, Phillips KM, Elnitsky C, Andresen EM. Validity of center for epidemiologic studies depression (CES-D) scale in a sample of Iraq and Afghanistan veteran. SAGE Open Med 2016;4:2050312116643906 [FREE Full text] [doi: 10.1177/2050312116643906] [Medline: 27127628]
24. Tapia AJ, Wagner F, Heredia MER, González-Forteza C. Study of depression in students from Mexico City and the state of Michoacán using the revised version of the CES-D. Salud Mental 2015;38(2):103-107. [doi: 10.17711/SM.0185-3325.2015.014]
25. Lacasse JJ, Forgeard MJ, Jayawickreme N, Jayawickreme E. The factor structure of the CES-D in a sample of Rwandan genocide survivors. Soc Psychiatry Psychiatr Epidemiol 2014;49(3):459-465. [Medline: 24173407]
26. Hann D, Winter K, Jacobsen P. Measurement of depressive symptoms in cancer patients: evaluation of the center for epidemiological studies depression scale (CES-D). J Psychosom Res 1999 May;46(5):437-443. [Medline: 10404478]
27. Boey KW. Cross-validation of a short form of the CES-D in Chinese elderly. Int J Geriatr Psychiatry 1999;14(8):608-617. [Medline: 10489651]
28. Vilagut G, Forero CG, Barbaglia G, Alonso J. Screening for depression in the general population with the center for epidemiologic studies depression (CES-D): a systematic review with meta-analysis. PLoS One 2016;11(5):e0155431. [Medline: 27182821]
29. Choi SW, Schalet B, Cook KF, Cella D. Establishing a common metric for depressive symptoms: linking the BDI-II, CES-D, and PHQ-9 to PROMIS depression. Psychol Assess 2014;26(2):513-527 [FREE Full text] [Medline: 24548149]
30. Cronbach LJ. Coefficient alpha and the internal structure of tests. Psychometrika 1951;16(3):297-334 [FREE Full text] [doi: 10.1007/BF02310555]
31. Bollen KA. A new incremental fit index for general structural equation models. Sociol Methods Res 1989;17(3):303-315. [doi: 10.1177/0049124189017003004]
32. Bentler PM. Comparative fit indexes in structural models. Psychol Bull 1990;107(2):238-246. [Medline: 2320703]
33. Browne MW, Cudek R. Alternative ways of assessing model fit. Sociol Methods Res 1993;21(2):230-258. [doi: 10.1177/0049124192021002005]
34. Cameron IM, Scott NW, Adler M, Reid IC. A comparison of three methods of assessing differential item functioning (DIF) in the hospital anxiety depression scale: ordinal logistic regression, Rasch analysis and the Mantel chi-square procedure. Qual Life Res 2014;23(10):2883-2888 [FREE Full text] [Medline: 24848597]
35. Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. Struct Equ Model a 1999;6(1):1-55. [doi: 10.1080/10705519909540118]
36. Sijtsma K, Emons WH, Bouwmeester S, Nyklíček I, Roorda LD. Nonparametric IRT analysis of quality-of-life scales and its application to the world health organization quality-of-life scale (WHOQOL-Bref). Qual Life Res 2008;17(2):275-290. [Medline: 18246447]
37. Stochl J, Jones PB, Croudace TJ. Mokken scale analysis of mental health and well-being questionnaire item responses: a non-parametric IRT method in empirical research for applied health researchers. BMC Med Res Methodol 2012 Jun 11;12:74 [FREE Full text] [doi: 10.1186/1471-2288-12-74] [Medline: 22686586]
38. Samejima F. Psychometricsociety. 1969. Estimation of latent ability using a response pattern of graded scores URL: https://www.psychometricsociety.org/sites/default/files/pdf/MN17.pdf [accessed 2017-08-11] [WebCite Cache ID 6sd8xdEVx]
39. Forero CG, Maydeu-Olivares A. Estimation of IRT graded response models: limited versus full information methods. Psychol Methods 2009 Sep;14(3):275-299. [doi: 10.1037/a0015825] [Medline: 19719362]
40. Baker FB. The basics of item response theory. College Park, MD: ERIC clearinghouse on assessment and evaluation; 2001.
41. Reeve BB, Hays RD, Bjorner JB, Cook KF, Crane PK, Teresi J, PROMIS Cooperative Group. Psychometric evaluation and calibration of health-related quality of life item banks: plans for the patient-reported outcomes measurement information system (PROMIS). Med Care 2007 May;45(5 Suppl 1):S22-S31. [doi: 10.1097/01.mlr.0000250483.85507.04] [Medline: 17443115]
42. Yen WM. Scaling performance assessments: strategies for managing local item dependence. J Educ Meas 1993;30(3):187-213 [FREE Full text] [doi: 10.1111/j.1745-3984.1993.tb00423.x]
43. Choi SW, Gibbons LE, Crane PK. Lordif: an R package for detecting differential item functioning using iterative hybrid ordinal logistic regression/item response theory and Monte Carlo. J Stat Softw 2011;39(8):1-30. [Medline: 21572908]

44. Crane PK, Gibbons LE, Jolley L, van BG. Differential item functioning analysis with ordinal logistic regression techniques. DIFdetect and difwithpar. Med Care 2006 Nov;44(11 Suppl 3):S115-S123. [doi: 10.1097/01.mlr.0000245183.28384.ed] [Medline: 17060818]
45. Zumbo BD. A handbook on the theory and methods of differential item functioning (DIF). Ottawa, Ontario, Canada: Directorate of Human Resources Research and Evaluation, Department of National Defense; 1999.
46. Crane PK, Gibbons LE, Ocepek-Welikson K, Cook K, Cella D, Narasimhalu K, et al. A comparison of three sets of criteria for determining the presence of differential item functioning using ordinal logistic regression. Qual Life Res 2007;16(1):69-84. [Medline: 17554640]
47. Cook KF, Bombardier CH, Bamer AM, Choi SW, Kroenke K, Fann JR. Do somatic and cognitive symptoms of traumatic brain injury confound depression screening? Arch Phys Med Rehabil 2011 May;92(5):818-823 [FREE Full text] [doi: 10.1016/j.apmr.2010.12.008] [Medline: 21530731]
48. Baylor C, Yorkston K, Eadie T, Kim J, Chung H, Amtmann D. The communicative participation item bank (CPIB): item bank calibration and development of a disorder-generic short form. J Speech Lang Hear Res 2013 Aug;56(4):1190-1208 [FREE Full text] [doi: 10.1044/1092-4388(2012/12-0140)] [Medline: 23816661]
49. Lord FM. Applications of item response theory to practical testing problems. Mahwah, NJ: Erlbaum Associates; 1980.
50. Veerkamp WJ, Berger MP. Some new item selection criteria for adaptive testing. J Educ Behav Stat 1997;22(2):203-226. [doi: 10.3102/10769986022002203]
51. Birnbaum A. Some latent trait models and their use in inferring an examinee’s ability. In: Statistical theories of mental test scores. Reading, MA: Addison-Wesley Pub Co; 1968:397-479.
52. Weiss DJ. Better data from better measurements using computerized adaptive testing. JMM 2011;2(1):1-27. [doi: 10.2458/v2i1.12351]
53. Walter OB. Adaptive tests for measuring anxiety and depression. In: Elements of adaptive testing. New York: Springer; 2009:123-136.
54. R Core Team. R-project. 2016. R: a language and environment for statistical computing URL: http://www.r-project.org/ [accessed 2017-08-11] [WebCite Cache ID 6sdFLbOk1]
55. Rosseel Y. Lavaan: an R package for structural equation modeling. J Stat Softw 2012;48(2):1-36 [FREE Full text] [doi: 10.18637/jss.v048.i02]
56. Van der Ark LA. New developments in Mokken scale analysis in R. J Stat Softw 2012;48(5):1-27 [FREE Full text] [doi: 10.18637/jss.v048.i05]
57. Chalmers RP. mirt: A multidimensional item response theory package for the R environment. J Stat Softw 2012;48(6):1-29 [FREE Full text] [doi: 10.18637/jss.v048.i06]
58. Mokken RJ. A theory and procedure of scale analysis with applications in political research. Berlin: Mouton; 1971.
59. Sijtsma K, Molenaar IW. Introduction to Nonparametric Item Response Theory. London: Sage Publications; 2002.
60. Maydeu-Olivares A, Joe H. Limited information goodness-of-fit testing in multidimensional contingency tables. Psychometrika 2006;71(4):713-732.
61. Forkmann T, Boecker M, Norra C, Eberle N, Kircher T, Schauerte P, et al. Development of an item bank for the assessment of depression in persons with mental illnesses and physical diseases using Rasch analysis. Rehabil Psychol 2009 May;54(2):186-197. [doi: 10.1037/a0015612] [Medline: 19469609]
62. Schalet BD, Pilkonis PA, Yu L, Dodds N, Johnston KL, Yount S, et al. Clinical validity of PROMIS depression, anxiety, and anger across diverse clinical samples. J Clin Epidemiol 2016;73:119-127. [Medline: 26931289]
63. Williams CD, Taylor TR, Makambi K, Harrell J, Palmer JR, Rosenberg L, et al. CES-D four-factor structure is confirmed, but not invariant, in a large cohort of African American women. Psychiatry Res 2007 Mar 30;150(2):173-180. [doi: 10.1016/j.psychres.2006.02.007] [Medline: 17291596]
64. Boisvert JA, McCreary DR, Wright KD, Asmundson GJ. Factorial validity of the center for epidemiologic studies-depression (CES-D) scale in military peacekeepers. Depress Anxiety 2003;17(1):19-25. [doi: 10.1002/da.10080] [Medline: 12577274]
65. Scalise K, Allen DD. Use of open-source software for adaptive measurement: Concerto as an R-based computer adaptive development and delivery platform. Br J Math Stat Psychol 2015;68(3):478-496 [FREE Full text] [Medline: 26061260]

Abbreviations

CAT: Computerized Adaptive Test
CES-D: Center for Epidemiologic Studies Depression Scale
CFA: Confirmatory factor analysis
CFI: Comparative fit index
DIF: Differential item functioning
GRM: graded response model
IRT: item response theory
RMSEA: root-mean-square error of approximation
TLI: Tucker Lewis Index

Edited by J Torous; submitted 06.02.17; peer-reviewed by J Apolinário-Hagen, T Donker; comments to author 12.04.17; revised version received 22.05.17; accepted 29.05.17; published 20.09.17.

Please cite as:
Loe BS, Stillwell D, Gibbons C
Computerized Adaptive Testing Provides Reliable and Efficient Depression Measurement Using the CES-D Scale
J Med Internet Res 2017;19(9):e302
URL: http://www.jmir.org/2017/9/e302/
doi:10.2196/jmir.7453
PMID:28931496

Original Paper

How the “Understanding Research Evidence” Web-Based Video Series From the National Collaborating Centre for Methods and Tools Contributes to Public Health Capacity to Practice Evidence-Informed Decision Making: Mixed-Methods Evaluation

Linda Chan[1], MPH; Jeannie Mackintosh[1], BA; Maureen Dobbins[1], BScN, PhD

National Collaborating Centre for Methods and Tools, School of Nursing, McMaster University, Hamilton, ON, Canada

Corresponding Author:
Maureen Dobbins, BScN, PhD
National Collaborating Centre for Methods and Tools
School of Nursing
McMaster University
175 Longwood Road South Suite 210a
Hamilton, ON,
Canada
Phone: 1 905 525 9140 ext 20450
Fax: 1 905 529 4184
Email: dobbinsm@mcmaster.ca

Abstract

Background: The National Collaborating Centre for Methods and Tools (NCCMT) offers workshops and webinars to build public health capacity for evidence-informed decision-making. Despite positive feedback for NCCMT workshops and resources, NCCMT users found key terms used in research papers difficult to understand. The Understanding Research Evidence (URE) videos use plain language, cartoon visuals, and public health examples to explain complex research concepts. The videos are posted on the NCCMT website and YouTube channel.

Objective: The first four videos in the URE web-based video series, which explained odds ratios (ORs), confidence intervals (CIs), clinical significance, and forest plots, were evaluated. The evaluation examined how the videos affected public health professionals’ practice. A mixed-methods approach was used to examine the delivery mode and the content of the videos. Specifically, the evaluation explored (1) whether the videos were effective at increasing knowledge on the four video topics, (2) whether public health professionals were satisfied with the videos, and (3) how public health professionals applied the knowledge gained from the videos in their work.

Methods: A three-part evaluation was conducted to determine the effectiveness of the first four URE videos. The evaluation included a Web-based survey, telephone interviews, and pretest and posttests, which evaluated public health professionals’ experience with the videos and how the videos affected their public health work. Participants were invited to participate in this evaluation through various open access, public health email lists, through informational flyers and posters at the Canadian Public Health Association (CPHA) conference, and through targeted recruitment to NCCMT’s network.

Results: In the Web-based surveys (n=46), participants achieved higher scores on the knowledge assessment questions from watching the OR (P=.04), CI (P=.04), and clinical significance (P=.05) videos but not the forest plot (P=.12) video, as compared with participants who had not watched the videos. The pretest and posttest (n=124) demonstrated that participants had a better understanding of forest plots (P<.001) and CIs (P<.001) after watching the videos. Due to small sample size numbers, there were insufficient pretest and posttest data to conduct meaningful analyses on the clinical significance and OR videos. Telephone interview participants (n=18) thought the videos’ use of animation, narration, and plain language was appropriate for people with different levels of understanding and learning styles. Participants felt that by increasing their understanding of research evidence, they could develop better interventions and design evaluations to measure the impact of public health initiatives.

Conclusions: Overall, the results of the evaluation showed that watching the videos resulted in an increase in knowledge, and participants had an overall positive experience with the URE videos. With increased competence in using the best available evidence, professionals are empowered to contribute to decisions that can improve health outcomes of communities.

(J Med Internet Res 2017;19(9):e286) doi:10.2196/jmir.6958

KEYWORDS

public health; public health practice; evidence-based practice; capacity building; continuing education; computer-assisted instruction

Introduction

Evidence-informed decision making in public health "integrates science-based interventions with community preferences to improve the health of populations" [1]. The use of evidence can impact intervention effectiveness in public health to minimize inequities in the community [2]. To cut through the complexity of applying evidence in practice, a clear method of how to do so is needed [3].

The National Collaborating Centre for Methods and Tools (NCCMT) has developed a seven-step process to engage with evidence-informed decision making in public health. The process outlines how to use the "best available evidence from research, context, and experience to inform and improve public health policy and practice" [4]. The seven steps of the evidence-informed public health process are as follows: defining the problem, searching for evidence, appraising the evidence, synthesizing the results of the evidence, adapting the evidence to the local context, implementing the knowledge translation strategy and the intervention, and evaluating the knowledge translation strategy and the outcomes of the intervention [5]. This process guides public health professionals through the process of finding and using the best available evidence when designing a public health intervention for their local context [4]. The NCCMT delivers training workshops on the evidence-informed public health process and critical appraisal to public health professionals across Canada. Through delivering these workshops, NCCMT staff identified understanding key concepts in research literature as a barrier to engaging with evidence.

Research literacy skills are necessary to facilitate the practice of evidence-informed decision-making [6]. Along with the lack of time to find and implement evidence, the lack of confidence in interpreting research is another barrier to using evidence in practice [7-9]. Public health professionals need to have dedicated time and access to high quality continuing professional development education on technical research concepts. This training will ensure that public health professionals understand research evidence to be able to engage in evidence-informed decision making [8,10].

NCCMT's Understanding Research Evidence (URE) video series was created to provide training on common research concepts. The URE videos were produced to be engaging and easily accessible for public health professionals. These videos are a Web-based resource that supports the practice of evidence-informed decision making in public health by building competency in interpreting research evidence. The topics in the series were chosen, based on feedback from previous workshop participants and the experience of NCCMT staff members. Communication, video production, and animation professionals were consulted in the development of the script and the creation of the videos. The first step in developing each video was writing a script. The initial draft was edited and revised substantially by JM to translate technical explanations of the concepts into plain language, while maintaining accuracy and clarity. MD has expertise on the technical aspects of the research concepts and has extensive experience training others to understand and conduct research. MD also reviewed the initial drafts to ensure the quality of the information. Throughout the editing process, the NCCMT team met regularly with a video production team and a cartoonist to discuss how best to convey the concepts visually. Narration of the final script was recorded on camera, and the cartoons were inserted. All videos in the series were produced in both English and French. The videos are available on the NCCMT website [11] and on YouTube. The first four URE videos in the series are on odds ratio (OR), clinical significance, confidence intervals (CIs), and forest plots. The videos describe the meaning of the concept and how to interpret or calculate the results. The first four videos in the URE series ranged from 3.5 min to 5.5 min.

An evaluation was conducted to examine public health professionals' experience with the first four videos in the URE video series to explain how the videos affected their work. In evaluations of education interventions, it is important to explore new ways to understand concepts, to justify the education intervention, and to improve the delivery of the intervention [12]. This evaluation examined both the delivery mode and the content of the videos. Specifically, this evaluation explored (1) whether the videos were effective at increasing knowledge on the four video topics, (2) whether public health professionals were satisfied with the videos, and (3) how public health professionals applied the knowledge gained from the videos in their work.

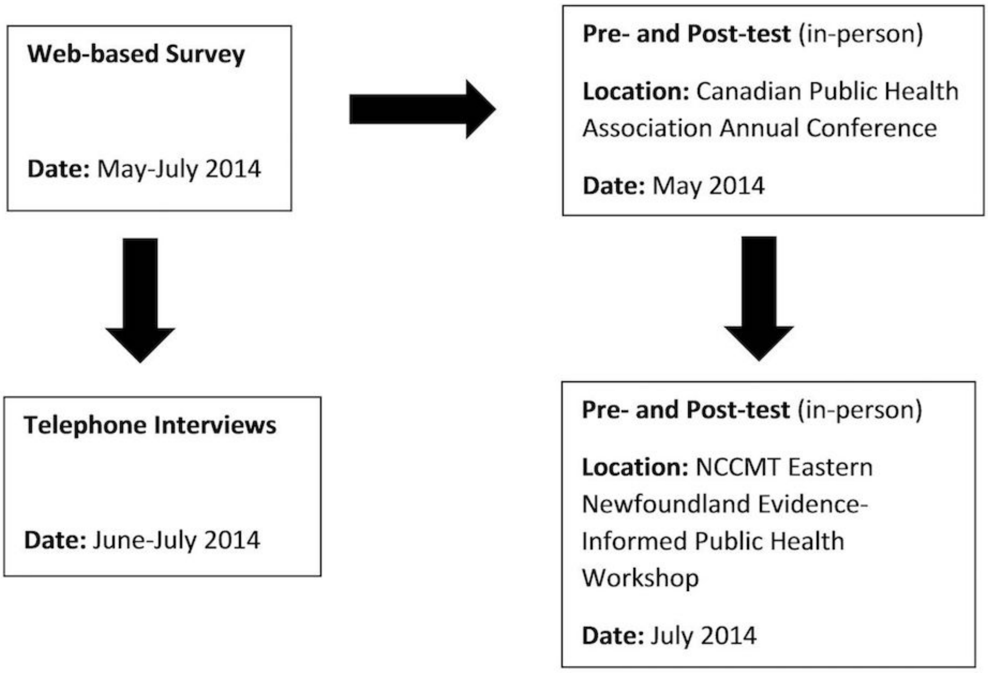
Methods

A mixed-methods approach was used for this evaluation. A Web-based survey, pretest and posttests, and semistructured telephone interviews were conducted to capture public health professionals' experiences with the URE videos. The data collected provided information on how to strengthen the design and delivery of the videos to maximize their utility for public health professionals. Figure 1 outlines the timeline of data collection. Both the Web-based survey and the pretest and posttests were used to collect data on the videos' ability to increase participants' knowledge on ORs, clinical significance, forest plots, and CIs. The Web-based survey increased the reach of this evaluation to public health professionals across Canada because of its easy accessibility. To build on the data collected from the Web-based survey, the pretest and posttests provided data using a stronger measure to assess the knowledge gained from the videos. Data from the telephone interviews complemented data from the Web-based survey and the pretest and posttests by exploring how the videos helped participants learn about research concepts and how participants used the videos in their public health work. Research ethics approval for this evaluation was obtained from the University of Toronto.

Sample

Since the NCCMT resources are intended for public health professionals across Canada, the evaluation sample included a broad range of Canadian public health professionals. Individuals were required to have watched at least one of the videos on ORs, clinical significance, CIs, and forest plots to participate in the Web-based survey and the semistructured telephone interviews. It was assumed that individuals would have access to computers and the Internet, and have basic computer skills, because most public health professionals are required to use computers as part of their role. Participation in the Web-based survey, pretest and posttest, and telephone interviews was voluntary.

Figure 1. Evaluation of the Understanding Research Evidence videos: data collection.



Data Collection

Web-Based Survey

Invitations to participate in the Web-based survey were sent out to four different public health Web-based email lists: the NCCMT Weekly Round-Up newsletter, the Health Evidence newsletter, Community Health Networking-Works newsletter, and the Public Health Knowledge Translation Network mailing list. Medical Officers of Health across Canada were also asked to invite their staff to participate in the Web-based survey through email. Public health professionals, who self-determined that they met sample criteria, could access the open survey through a link provided in the invitation. The Web-based survey was conducted from May 28, 2014 to July 2, 2014 on Qualtrics Research Suite (Qualtrics, Salt Lake City, Utah), and it collected data on who had been using the videos, how the videos were being accessed, and assessed participants' knowledge on ORs, CIs, forest plots, and clinical significance. A letter of information introducing the survey and the purpose of the evaluation was included in the first page of the Web-based survey. Consent to participate in the survey was provided from participants when they clicked to continue past the first page. The data collected was anonymous, and all the responses were stored on an encrypted computer. The multiple-choice knowledge assessment questions tested the participants' ability to understand, interpret, and apply the research concepts (see Multimedia Appendix 1). The knowledge assessment questions were written by LC and vetted by MD, who has expertise on the concepts through her research and teaching commitments. The response options provided different explanations and interpretations of the research concept with only one correct answer. In order to demonstrate that they had knowledge of the concept, participants' needed to select the correct answer. Web-based survey participants had the option to review and change their answers throughout the survey. At the end of the survey, participants had the option to enter their name for a draw to win a $50 gift card (CAD) to a Canadian bookstore.

Pretest and Posttest

Two sets of in-person pretest and posttests were conducted concurrent to the Web-based survey. The pretest and posttests assessed whether knowledge was gained from watching the URE videos. Four separate tests were created for each of the URE videos. The knowledge assessment questions were written in the same manner as the knowledge assessment questions in the Web-based survey (see Multimedia Appendix 1). Each test had a set of 5 to 6 multiple-choice questions that asked about key ideas necessary to understand the research concept.

The first pretest and posttest was conducted in-person at the Canadian Public Health Association (CPHA) annual conference in May 2014. Participants were recruited using informational flyers and posters at the NCCMT's CPHA exhibition booth. A message inviting individuals to participate in a video-viewing session was also included in the NCCMT Weekly Round-Up

E-newsletter. A random schedule was created for the video showings. Twelve video-viewing sessions were held for the four URE videos. Participants were blinded to this schedule, so they did not know which of the four URE videos was being shown at the start of their participation. At the video-viewing sessions, individuals were invited to complete an informed consent form, a pretest, watch one of the four URE videos, and complete a posttest.

Due to the low participation rate for the initial pretest and posttest session, another session was conducted using iClicker technology at the NCCMT's Eastern Newfoundland evidence-informed public health and critical appraisal workshops in July 2014. Additional data was collected on the forest plot and CI videos.

Telephone Interview

The semistructured telephone interviews explored the following: why public health professionals watched the URE videos, whether the videos were engaging, how the videos facilitated learning, and how the knowledge gained from the videos was used in participants' professional work. Individuals who completed the Web-based survey were invited to participate in an interview. However, this recruitment method only rendered 2 participants. To recruit additional interview participants, individuals from the NCCMT network were contacted and invited to participate in an interview. Once individuals agreed to participate in a telephone interview, any individuals who had not seen one of the URE videos were asked to watch one of the four videos prior to their scheduled interview time.

To ensure that participants provided informed consent to participate in the interview, they were emailed a letter of information about the evaluation prior to the interview. At the beginning of the interview, a brief overview of the evaluation was provided to the participants to ensure that they understood for what they were providing consent. After the overview, participants were asked to provide verbal consent to participate in the interview and to have the interview recorded.

Data Analysis

Web-Based Survey

Descriptive frequencies were used to analyze demographic data. Participants were given a point for the answers they got correct in the knowledge assessment questions in the survey. A three by two chi-square (χ^2) test was used to analyze the relationship between the videos watched and the total knowledge assessment score. All responses were reviewed to determine the completeness of each survey. Internet Protocol addresses were also reviewed to identify any potential duplicate entries from the same user. Incomplete questionnaires were included in analysis. Missing data were reported in the results section.

Pretest and Posttest

A paired *t*-test was used to analyze the amount of knowledge gained on each research concept before and after watching the video shown for the pretest and posttest. The Web-based survey data and pretest and posttest data were analyzed using SPSS Statistics (IBM, Chicago, Illinois).

Telephone Interviews

A framework analysis approach was used to analyze the telephone interview data. Interviews were transcribed and coded by LC to gain familiarity with the data. The interview transcripts were reviewed initially to identify themes about participants' thoughts and experiences that emerged with the URE videos. The interviews were indexed iteratively by systematically reapplying codes that emerged from one interview to the entire data set. Finally, relationships between themes were explored to make connections and associations between concepts [13]. Dedoose software (SocioCultural Research Consultants, Los Angeles, California) was used for analysis of the interviews.

Results

Web-Based Survey: Demographic Information

Results from the Web-based survey (n=46) described the type of public health professionals who had watched the URE videos. The survey respondents spanned across Canada from British Columbia, Alberta, Saskatchewan, Manitoba, Ontario, Newfoundland, Nova Scotia, Yukon, and the Northwest Territories. Table 1 summarizes the survey respondents' education level, amount of public health work experience, and occupation roles. Table 2 describes how respondents found out about the URE videos and where they watched the videos.

Web-Based Survey: Comprehension of the Research Concepts

In the Web-based survey, 33% (15/46) of participants stated they had a good understanding of research concepts, and 96% (44/46) of survey participants felt that they gained knowledge from watching the videos. Despite many participants having a baseline understanding of the research concepts, results from the chi-square test showed there was an increase in the number of knowledge assessment questions that participants got right from watching the URE videos.

Table 3 summarizes the results of the Web-based survey knowledge assessment questions. For the OR, CI, and clinical significance videos, a relationship was found between having watched the URE video and a higher score on the knowledge assessment questions (OR χ^2_2 [N=46]=6.7, P=.04; CI χ^2_2 [N=46]=6.7, P=.04; clinical significance χ^2_2 [N=46]=6.0, P=.05). The results for the forest plot video were not statistically significant (forest plots χ^2_2 [N=46]=4.3, P=.12; P<.05). There was insufficient data to determine whether or not watching the forest plot video resulted in higher scores on the knowledge assessment questions. For the other three videos, the Pearson correlation ([OR: r=.348], [Clinical Significance: r=.286], [95% CI: r=.348]) indicates that there was a medium correlation between having watched the URE video and participants' score on the knowledge assessment questions.

Table 1. Location, education, public health work experience, and occupation of Web-based survey participants.

Demographic	Number of participants (%)
Province[a]	
British Columbia	2 (4)
Alberta	4 (8)
Saskatchewan	2 (4)
Manitoba	1 (2)
Ontario	23 (51)
Newfoundland and Labrador	10 (22)
Nova Scotia	1 (2)
Yukon	1 (2)
Northwest Territories	1 (2)
Total	45
Education	
Bachelor’s degree	23 (50)
Master’s degree	20 (43)
MD (Medical Doctor)	1 (2)
Doctorate (PhD, EdD)	2 (4)
Total	46
Public health work experience[a]	
Less than 1 year	7 (15)
1-5 years	8 (17)
6-10 years	9 (20)
11-15 years	11 (24)
16-20 years	2 (4)
21-25 years	3 (6)
25+ years	5 (11)
Total	45
Occupation[a]	
Health promotion/Educator	8 (17)
Dietitian/Public health nutritionist	4 (8)
Public health nurse/Registered nurse	10 (22)
Academic/Professor	2 (4)
Program Evaluator	3 (6)
Consultant/Specialist	2 (4)
Researcher/Research analyst	6 (13)
Other	10 (22)
Total	45

[a] Number of participants does not equal 46 because of missing data.

Table 2. Accessing the URE videos.

Access to the videos	Number of participants (%)
Learned about videos[a]	
NCCMT weekly round-up newsletter	2 (4)
NCCMT website	4 (8)
NCCMT workshops	2 (4)
Twitter	1 (2)
Canadian Public Health Association Conference 2014	23 (51)
Word of mouth	10 (22)
Total	1 (2)
Location where videos were watched	
NCCMT website	35 (76)
YouTube	3 (6)
Training session	6 (13)
Not sure	2 (4)
Total	46

[a] Number of participants does not equal 46 because of missing data.

Table 3. Frequency tables of Web-based survey knowledge assessment scores.

Video	Knowledge assessment score	Did not watch[a]	Watched[a]	Total	*r* (*P* value)
Odds ratio video					.348 (.04)
	0	4	2	6	
	1	6	18	24	
	2	2	14	16	
	Total	12	34	46	
Forest plot video					.296 (.12)
	0	4	4	8	
	1	4	12	16	
	2	3	19	22	
	Total	11	35	46	
Clinical significance video					.286 (.05)
	0	6	3	9	
	1	8	26	34	
	2	1	2	3	
	Total	15	31	46	
Confidence interval video					.348 (.04)
	0	4	2	6	
	1	6	18	24	
	2	2	14	16	
	Total	12	34	46	

[a] A correction factor was included in the χ^2 analysis on SPSS to account for the small cell sizes in the frequency table.

Table 4. Pre- and posttest knowledge assessment scores.

Video	Mean difference	df[a]	Paired *t*-test	*P* value	Cohen *d*	Standard Deviation	N
Odds ratio	26.67	2	4.00	.057	2.309	11.55	3
Clinical significance	13.33	5	1.35	.235	0.551	25.30	6
Forest plots	30.49	60	5.71	.000	0.731	32.17	61
Confidence interval	38.36	54	6.84	.000	0.922	26.00	55

[a] df: degrees of freedom.

Pretest and Posttest: Comprehension of the Research Concepts

The pretest and posttests had a total n of 124. Eleven individuals participated in the first video-viewing session at CPHA. There were 2 participants for the CI video, 6 participants for the clinical significance video, 1 participant for the forest plot video, and 3 participants for the OR video. From the Eastern Newfoundland workshop, there were an additional 60 participants for the forest plot video, and an additional 53 participants for the CI video.

Pretest and posttest results are summarized in Table 4. Results from paired *t*-tests showed there was an increase in knowledge on forest plots and CIs after watching the URE videos. A statistically significant difference was found for the forest plots' and the CIs' pretest and posttest data ([Forest plots: t_{61}=5.710, $P<.001$, d=0.731], [CIs: t_{55}=6.835, $P<.001$, d=0.922]). There was a medium-large effect size for the forest plot video pretest and posttest, and a large effect size for the CI video pretest and posttest. Moreover, the OR video pretest and posttest (d=2.309) had a very large effect size, and the clinical significance video pretest and posttest (d=0.551) had a medium effect size. However, statistical significance was not detected for the clinical significance (n=6) and OR (n=3) videos due to small sample sizes.

Telephone Interviews

Most telephone interview participants (n=18) had watched multiple videos: 15 participants watched the CI video; 12 watched the OR video; 13 watched the clinical significance video; and 12 watched the forest plot video. A summary of the telephone interview results can be found in Textboxes 1-3.

Comprehension of Research Concepts

In the telephone interviews, participants stated that the videos acted as a refresher to reinforce concepts they had learned in their Masters' programs, statistic courses, and continuing education courses. The videos reminded participants of the important points of a research concept. Participants indicated the videos helped them understand how to analyze data to determine whether they were looking at good evidence. One participant noted:

> *It's a refresher but it also reminds people how doable it really is. I just thought after seeing it, I thought, this actually makes somebody feel like they could actually just take a pen and paper and do some simple ratios.*

Textbox 1. Summary of telephone interview theme 1: increased comprehension of research concepts.

Main findings:

- URE videos helped participants understand common concepts in research evidence.
- URE videos reinforced important points to know about research concepts.

Quotations:

- "There was a general increase in knowledge and understanding, but I think more importantly for me, there was a reinforcement of what I already knew." (Participant 1)
- "They may have heard the term, but it's not something that they use everyday, so we're just trying to build that capacity within our organization." (Participant 6)
- "It is a nice refresher. It (URE videos) doesn't necessarily tell you anything that you didn't already know, but it reminds you of the important points that sometimes get lost." (Participant 11)

Textbox 2. Summary of telephone interview theme 2: learning through the videos.

Subtheme: Delivery of the content

- Main findings
  - Narration and animation in the videos acted as learning aids that supported both visual and auditory learners.
  - URE videos were accessible because they avoided technical jargon.
  - The short length allowed public health professionals to use the videos during their workday.
  - Videos were used both individually and in groups.
- Quotations
  - “I think it (URE Videos) really speaks to all adult learners, you know the ones that like to see things, the ones that like to hear things. I think it speaks to a varied amount of learners.” (Participant 9)
  - “I found that the language that was used was quite clear and plain. The presenter was able to simplify what might be thought of as complex.” (Participant 18)
  - “The length was good. They (URE Videos) weren’t too long, but they were long enough to explain the concept.” (Participant 5)
  - “I can also see these being used at a team meeting. I don’t know if this has been done, but you could pull this (URE Videos) up at a team meeting and say let’s spend the next seven team meetings just do one a week and go over it. What a great way to keep your learning up to date.” (Participant 10)

Subtheme: Content of the videos

- Main findings
  - Examples in the videos were relevant to public health and helped illustrate the key ideas of the research concept.
  - Videos focused on how to interpret a research concept.
- Quotations
  - “I liked how the videos related some of the concepts to practical examples in public health. Giving public health examples we could all relate to made it all the more easier to understand the concept.” (Participant 3)
  - “Some videos I found were more for professionals or decision-makers who want to learn how to interpret some statistical concepts, but some other videos I thought were giving some more information for professionals who would reproduce or would have to make that kind of statistical method…” (Participant 16)
  - “So what’s in it for me is whether a relationship exists, and what the strength of that relationship is between the intervention and the control.” (Participant 14)

Textbox 3. Summary of telephone interview theme 3: supporting the use of research evidence.

Main findings

- The URE videos helped public health professionals develop core competencies for evidence-informed public health.
- Understanding common research concepts helped public health professionals assess the quality and strength of the research.
- The videos provided a common language to talk about intervention effectiveness.

Quotations

- “One of my main tasks in my job is to make sure that the staff at the health unit have the skills to do evidence-informed public health. So you know, I’m always looking for really good resources that I can pass on to the staff to increase their skill, probably in the appraising stage of articles.” (Participant 8)
- “I got a much better sense of, like if I’m reading a research report or kind of report that is using statistics, I will have a much better sense of what it means and hopefully be able to use it in application.” (Participant 12)
- “It (URE videos) gave me language also to explain it to others in my workplace… I found the videos gave me some clear examples and language that I could help explain it to others.” (Participant 3)

Learning Through the URE Videos

Delivery of the Content

The telephone interviews revealed that narration and animation in the videos provided learning aids to understand the research concepts covered. One participant thought the videos spoke to all adult learners. The "words, numbers, and pictures" used in the videos appealed to both visual and auditory learners. The narration in the videos helped participants engage with the video content with minimal effort. Participants could just "sit back and listen." Participants found that the use of clear and plain language helped to explain the technicalities of the research concept, and gave them a clear understanding of the topic. They also found that avoiding technical jargon simplified complicated concepts, making the videos easy to understand. Participants found the videos enjoyable to listen to, and the concepts were accessible to those who were not familiar with research. Participants also thought the narration in the videos made the content easier to understand for auditory learners. They found repeating definitions of the concepts using similar language and having the narrator speak slowly facilitated learning for auditory learners.

Participants, who were visual learners, appreciated the use of animations and visual diagrams. Moreover, participants felt the combination of narration with animations made it easier to follow the content. Participants thought the combination of narration and animation brought the examples used in the videos to life, and made technical research concepts less intimidating. Participants also thought the animations helped to "lighten the gravity of the information" and kept the videos dynamic, which helped maintain viewers' interest in the content. One participant noted:

> *The animations were good. They kind of kept you watching and kept you interested and intrigued. They helped you, at certain points, I guess, they helped you understand some of the concepts just viewing the animations, which is not possible sometimes with just having a person talk.*

Overall, participants found the videos appropriate and pleasant to watch. They found the videos short and concise, and liked that the videos were quick to watch. They thought that the videos were long enough to explain the concept, but short enough to accommodate interruptions during the workday. The short length of the videos also allowed participants to use the videos as a quick reference.

Participants also noted how the videos could be used individually and in groups to facilitate learning. They felt the videos could be used individually when interpreting the results of a study or in a group. Participants also described how the videos could be used for peer learning in a group setting such as a quick refresher before a meeting, in journal clubs, in classroom settings, or in a larger continuing education program for public health professionals. Most participants preferred watching the videos in a group setting. One participant explained how she used the videos in a meeting:

> *We reviewed the videos, and then discussed how we would present it to our teammates. What we decided to do was, we brought a laptop and projector into our team meeting, and we showed the videos to our teammates, and then what we did is we talked about it in terms of work that we are currently doing, and we were able to also bring something that we were actually reviewing. So for example, she brought a research article that had a forest plot in it, so we showed that to the team, and explained how this was a real article that was being reviewed. And what I did is, I had been working with our [epidemiology] department, and was currently reviewing health status data. So I brought one of the tables from a health status report that they had prepared for our team. We were able to talk about current and relevant uses and application to the work we were doing.*

Content of the Videos

The telephone interviews demonstrated that the examples included in the videos were relevant and common topics in public health, which helped participants understand and relate to the content. Participants felt that the examples illustrated what the concept meant and broke down the content into different parts. One participant stated:

> *By using the examples, it made the definition of odds ratio or whatever I was looking at more real to me, again versus reading it in a textbook. I probably wouldn't have had as decent of an understanding as I did by having the example, it just brought it more to life for me.*

Participants found that all the videos focused on explaining how to interpret a research concept, with some also explaining how to calculate or reproduce a research concept. For example, one participant felt that the video on CIs focused more on the interpretation of the concept, whereas the video on OR explained how to calculate an OR as well as how to interpret it. Participants thought that the videos focused on how to interpret a research concept and were helpful and simpler to follow.

Supporting the Use of Research Evidence—From Application to Practice

Participants thought that by increasing the use of research evidence in their practice, they could develop better program interventions and design evaluations to measure the impact of public health initiatives. Participants felt that the URE videos helped public health professionals develop a core competency for evidence–informed decision-making. They felt the videos were effective at building skills in technical research concepts, which allowed them to better engage with research evidence. They also felt that understanding research evidence allows public health professionals to learn about different types of intervention that exist for an issue and why the interventions are effective in particular contexts. One participant expressed:

> *I really am hoping that our team as we move forward will start to focus more on looking at the type of interventions we do, and making sure that they are research-based and evidence-based, and that would mean we would have to start reading the research more. And I think that's how this kind of information*

> *can help our team. When we're reading the research, knowing what a confidence interval is, what that means, that way we can better decide if an intervention is something that is worthy of us putting into place or not. I can see us definitely learning from these videos how to better understand the research that has already been done in our field.*

Furthermore, participants felt that the URE videos provided them with a deeper understanding of how to interpret and understand the details of research data. One participant noted how understanding common research concepts allowed public health professionals to assess the quality and strength of research evidence. Another participant thought the ability to interpret research evidence could also help public health professionals design evaluations to assess interventions. This participant felt that research and evaluation was important in keeping public health accountable for its action. Another participant stated that a clearer understanding of research concepts could make evaluation less intimidating.

Finally, participants found the videos introduced them to more approachable ways to discuss data by providing a common language to talk about intervention effectiveness. They felt the language to talk about intervention effectiveness would impact the uptake of research evidence in public health work. This participant stated:

> *It gave me language also to explain it to others in my workplace. We do work with, we do have some students working with us, and it's always helpful to have a different way to kind of explain a concept, than what you have used in the past. And I found the videos gave me some of those clear examples and language that I could help explain it to others.*

Discussion

Principal Findings

The Web-based survey and the pretest and posttest demonstrated that the URE videos were effective in increasing public health professionals' knowledge on ORs, clinical significance, CIs, and forest plots. The results of the survey and pretest and posttest were further supported by the themes uncovered from the telephone interviews. Findings from the telephone interviews demonstrated how the URE videos were an effective continuing professional development resource that builds public health capacity to use research evidence. Public health professionals who participated in this evaluation indicated that they learned how to interpret or understand the research concept covered in the video. They were satisfied with the delivery methods and content of the videos and felt they could apply what they had learned in their professional practice. The evaluation found a relationship between watching the URE video and the ability to recall and comprehend the concepts.

Public health professionals found the content of the URE videos applicable to their work. They thought the URE videos could help them with program implementation and evaluation, because an understanding of research concepts would make research findings and the idea of measurement more approachable. The URE videos also provided the public health professionals with language to talk about data, which increased discussion about research and led to increased engagement with research evidence. Overall, public health professionals were receptive of the URE videos and found the resource to be useful to their work.

The public health workforce is broad. It ranges from medical officers of health to biostatisticians, health promoters, and other positions. The ability to understand common research concepts is a foundational competency for the implementation of effective interventions. Although public health professionals in different roles may use research evidence for different purposes, it is important that all public health professionals have a baseline understanding of common research concepts and the language to meaningfully discuss research evidence with their colleagues. The URE videos can build public health professionals' capacity to engage with research evidence when planning, implementing, and evaluating public health interventions. This will ensure all public health interventions, from clinical interventions to programming, are effective at improving the health of individuals and communities.

Limitations

Since the URE videos were only made available on the Internet in 2013, their impact on the practices of public health professionals may not yet be fully known. Future evaluations should further explore the longitudinal impact of the URE videos. Moreover, findings from this evaluation may be limited due to bias from selection and self-assessment. The majority of participants in the Web-based survey were registered NCCMT users. The public health professionals who watched the URE videos may have already be interested and more competent in using research for evidence-informed public health than other public health professionals. However, public health professionals who had not seen the URE videos were asked to watch the videos in order to participate in the telephone interviews. Many of these public health professionals were not familiar with the NCCMT but still addressed the importance of understanding and using research evidence in public health practice. Additionally, self-assessment was used to gather data on participants' background knowledge on research concepts and whether or not they felt they had gained knowledge from watching the videos. Participants may not have accurately responded to these questions due to distorted self-perceptions.

Furthermore, the external and internal validity of the findings should be considered. The Web-based survey had a low response rate, which may affect the generalizability of the results. Although the survey demonstrated that a medium correlation exists between knowledge level (as determined by knowledge assessment scores) and the clinical significance and OR videos; due to small sample sizes, no statistical significance was detected for the pretest and posttest results for the two videos. Similarly, a medium-large effect size was detected in the pretest and posttest for the forest plot video; however, a statistical significant correlation between knowledge level and the forest plot video was not detected in the Web-based survey. Additional testing of the clinical significance, OR, and forest plot videos need to be conducted to gather more consistent results. Also, none of the knowledge assessment questions in the Web-based survey or pretest and posttest were piloted due to time restraints. It is unknown whether the questions adequately captured the constructs needed to understand the OR, clinical significance, CI, and forest plot concepts. Finally, Web-based survey and telephone participants had the option of watching the videos in French or English. However, data was not collected on the language used to watch the videos.

Future evaluation of the URE videos should include the following: test of the validity of the knowledge assessment questions, increase in sample sizes and the number of knowledge assessment questions asked for each video (to increase the statistical significance of the findings), and a determination of whether knowledge gained differed when watching the videos in French versus English. Despite the limitations of the evaluation, the findings provided useful insights into public health professionals' experience with the URE videos, and the manner in which they intend to use the videos in their practice.

Future Applications

Since public health is a broad multidisciplinary field, wide-reaching and targeted continuing education strategies are necessary to facilitate universal public health competencies [14,15]. Continuing professional development in public health aims to maintain core professional competencies for public health practitioners. The ability to interpret research evidence is a core competency for evidence-informed decision making [6,8,10]. Bridging knowledge gaps in understanding research concepts and interpreting research evidence can address barriers to evidence-informed decision making in public health.

Continuing professional development training provides the knowledge and skills necessary to improve population health [14]. Public health professionals are motivated to participate in continuing professional development activities due to peer pressure, to maintain a high standard of professional competence, and because they desire self-advancement [16]. However, educational strategies and technology used to deliver continuing professional development need to be considered to facilitate effective learning within the constraints of professional practice.

Web-based technology is becoming a popular way to administer continuing professional education in public health. According to a study conducted by the American Public Health Association [17], 76% of public health professionals prefer continuing education training to be delivered through webcasts and webinars, and 65% prefer Web-based, self-paced training. The Centers for Disease Control and Prevention in the United States (Public Health Training Network), the United Kingdom's Department of Health (Teaching Public Health Networks), and Canada's Public Health Agency of Canada (Skills Online) have all created Web-based continuing professional development training programs for public health professionals [10,18-20]. However, there are still a limited number of free, open-access learning resources on the Web for public health [21].

Effective Web-based continuing education programs should use appropriate delivery methods and consider the type of information being provided [19]. Public health professionals were satisfied with the delivery and content of the videos. They liked the URE videos because the combination of animation, narration, and plain language made the content clear. Enjoyment in learning can motivate learners and allow them to relax to better engage with the content [22]. Moreover, time is a major barrier to engaging in evidence-informed public health [7]. The videos were short enough to fit into existing work schedules.

According to Friedman [23], effective continuing education programs should appeal to a range of learners and consider the needs of individuals with different learning styles to support learners to be proactive in their learning. The videos were delivered using strategies that targeted different types of professional adult learners. The URE videos addressed auditory, visual, independent, and group learners. The fact that these videos were dynamic and included animation, examples, and narration may have facilitated learning of the concepts.

The URE video series add to the existing public health professional development resources by offering specific training on understanding common research concepts that are freely accessible on the Internet. Since the completion of this evaluation, additional videos have been created on the following topics: how to find research evidence, the evidence-informed decision making framework, types of reviews, relative risk, number needed to treat, *P* values, and the standardized mean difference [11]. Public health professionals who participated in this evaluation suggested additional topics they would like to see covered in the URE Web-based video series, which included tests for heterogeneity, statistical significance, and effect sizes. Participants also suggested that future videos could cover how to understand qualitative and mixed-methods research.

Conclusions

The URE video series is one of the many resources that the NCCMT offers to build capacity for evidence-informed public health. The lack of skill in understanding common research concepts was identified as a barrier in engaging with evidence-informed decision making. Public health professionals need to be able to understand research evidence to be able to use it in their practice.

Public health professionals' knowledge of common research concepts was impacted by the URE videos. The delivery and content of the URE videos helped to facilitate learning. An increased understanding of research concepts made public health professionals feel that they could use research evidence to inform program implementation, and made program evaluation more approachable; it also gave them the language to discuss data and research with their colleagues. Overall, the URE videos were received well, and public health professionals hoped for more videos on different research topics in the future. The NCCMT will continue to expand and develop the Web-based URE video series and other resources to build capacity in evidence-informed decision making among Canadian public health professionals, and empower them to use the best available evidence when making decisions to improve community and population health outcomes.

Acknowledgments

The authors thank Pamela Forsyth for her support with the implementation of this evaluation. The authors would also like to acknowledge Dr Suzanne Jackson and Dr Michael Goodstadt at the Dalla Lana School of Public Health, University of Toronto, for their contribution in advising on the design and analysis of this evaluation.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Web-based survey and pretest/posttest knowledge assessment questions.

[PDF File (Adobe PDF File), 321KB - jmir_v19i9e286_app1.pdf]

References

1. Kohatsu ND, Robinson JG, Torner JC. Evidence-based public health: an evolving concept. Am J Prev Med 2004 Dec;27(5):417-421. [doi: 10.1016/j.amepre.2004.07.019] [Medline: 15556743]
2. Rychetnik L, Wise M. Advocating evidence-based health promotion: reflections and a way forward. Health Promot Int 2004 Jun;19(2):247-257. [Medline: 15128716]
3. Tannahill A, Kelly MP. Layers of complexity in interpreting evidence on effectiveness. Public Health 2013 Feb;127(2):164-170 [FREE Full text] [doi: 10.1016/j.puhe.2012.11.011] [Medline: 23332513]
4. Ciliska D, Thomas H, Buffett C. NCCMT.: National Collaborating Centre for Methods and Tools; 2010. An introduction to evidence-informed public health and a compendium of critical appraisal tools for public health practice URL: http://www.nccmt.ca/pubs/2008_07_IntroEIPH_compendiumENG.pdf [accessed 2016-08-02] [WebCite Cache ID 6jS2n3xqn]
5. National Collaborating Centre for Methods and Tools. NCCMT. 2014. Evidence-informed public health URL: http://www.nccmt.ca/eiph/index-eng.html [accessed 2016-08-02] [WebCite Cache ID 6jS2j3eIE]
6. Porter J, McGrath R, Costello A. Development of an evidence-based data skills curriculum for public health professionals. Am J Health Educ 2008 Sep;39(5):314-317. [doi: 10.1080/19325037.2008.10599056]
7. Armstrong R, Pettman TL, Waters E. Shifting sands - from descriptions to solutions. Public Health 2014 Jun;128(6):525-532. [doi: 10.1016/j.puhe.2014.03.013] [Medline: 24916424]
8. Rutland JD, Smith AM. Information needs of the 'frontline' public health workforce. Public Health 2010 Nov;124(11):659-663. [doi: 10.1016/j.puhe.2010.06.002] [Medline: 20888017]
9. Oliver K, Innvar S, Lorenc T, Woodman J, Thomas J. A systematic review of barriers to and facilitators of the use of evidence by policymakers. BMC Health Serv Res 2014 Jan 03;14:2 [FREE Full text] [doi: 10.1186/1472-6963-14-2] [Medline: 24383766]
10. Pilkington P. Improving access to and provision of public health education and training in the UK. Public Health 2008 Oct;122(10):1047-1050. [doi: 10.1016/j.puhe.2008.03.008] [Medline: 18662819]
11. National Collaborating Centre for Methods and Tools. NCCMT. 2016. Multimedia URL: http://www.nccmt.ca/resources/multimedia [accessed 2016-08-02] [WebCite Cache ID 6jS2Ovb6h]
12. Ringsted C, Hodges B, Scherpbier A. 'The research compass': an introduction to research in medical education: AMEE Guide no. 56. Med Teach 2011;33(9):695-709. [doi: 10.3109/0142159X.2011.595436] [Medline: 21854147]
13. Green DG, Thorogood N, editors. Analysing Qualitative data. In: Qualitative Methods for Health Research (Introducing Qualitative Methods series). Thousand Oaks: Sage Publications Ltd; 2009:195-228.
14. Bennetts C, Elliston K, Maconachie M. Continuing professional development for public health: an andragogical approach. Public Health 2012 Jun;126(6):541-545. [doi: 10.1016/j.puhe.2012.01.032] [Medline: 22633081]
15. Balmer JT. Transforming continuing education across the health professions. J Contin Educ Nurs 2012 Aug;43(8):340-341. [doi: 10.3928/00220124-20120725-02] [Medline: 22849626]
16. Haywood H, Pain H, Ryan S, Adams J. Engagement with continuing professional development: development of a service model. J Allied Health 2012;41(2):83-89. [Medline: 22735821]
17. Neuberger JS, Montes JH, Woodhouse CL, Nazir N, Ferebee A. Continuing educational needs of APHA members within the professional public health workforce. J Public Health Manag Pract 2014;20(5):530-533. [doi: 10.1097/PHH.0b013e3182a9c0f5] [Medline: 24100241]
18. Orme J, Pilkington P, Gray S, Rao M. Teaching public health networks in England: an innovative approach to building public health capacity and capability. Public Health 2009 Dec;123(12):800-804. [doi: 10.1016/j.puhe.2009.10.008] [Medline: 19914668]
19. Jamison B. Chasing the technology rabbit. J Public Health Manag Pract 2001 Jul;7(4):50-54. [Medline: 11434040]
20. Public Health Agency of Canada. Phac-aspc. 2014. Skills online URL: http://www.phac-aspc.gc.ca/php-psp/ccph-cesp/index-eng.php [accessed 2016-08-02] [WebCite Cache ID 6jS2TkZbs]

21. Angell C, Hemingway A, Hartwell H. Surfing the net for public health resources. Public Health 2011 Aug;125(8):547-553. [doi: 10.1016/j.puhe.2011.02.007] [Medline: 21802101]
22. Prensky M. Fun, play, and games: what makes games engaging. In: Digital Game-Based Learning. St Paul: Paragon House Publishers; 2001.
23. Friedman B. Teaching and learning styles. In: How to Teach Effectively: A Brief Guide, 2nd Edition. Chicago: Lyceum Books; 2013:1-16.

Abbreviations

CPHA: Canadian Public Health Association.
NCCMT: National Collaborating Centre for Methods and Tools.
URE: Understanding Research Evidence.

Edited by CL Parra-Calderón; submitted 08.11.16; peer-reviewed by D Tao, H Robitaille; comments to author 03.01.17; revised version received 08.03.17; accepted 28.04.17; published 28.09.17.

Please cite as:
Chan L, Mackintosh J, Dobbins M
How the “Understanding Research Evidence” Web-Based Video Series From the National Collaborating Centre for Methods and Tools Contributes to Public Health Capacity to Practice Evidence-Informed Decision Making: Mixed-Methods Evaluation
J Med Internet Res 2017;19(9):e286
URL: http://www.jmir.org/2017/9/e286/
doi:10.2196/jmir.6958
PMID:28958986

Original Paper

Video Game Intervention for Sexual Risk Reduction in Minority Adolescents: Randomized Controlled Trial

Lynn E Fiellin[1,2,3,4], MD; Kimberly D Hieftje[1,3,4], PhD; Tyra M Pendergrass[1,3,4], MEM; Tassos C Kyriakides[5], PhD; Lindsay R Duncan[1,6], PhD; James D Dziura[5], PhD; Benjamin G Sawyer[1,7]; Linda Mayes[2,4], MD; Cindy A Crusto[3,4,8], PhD; Brian WC Forsyth[3,4], MBChB, FRCPC; David A Fiellin[1,3,4,5], MD

[1]play2PREVENT Lab, Yale School of Medicine, Yale University, New Haven, CT, United States
[2]Yale Child Study Center, New Haven, CT, United States
[3]Center for Interdisciplinary Research on AIDS, Yale School of Public Health, Yale University, New Haven, CT, United States
[4]Yale School of Medicine, Yale University, New Haven, CT, United States
[5]Yale School of Public Health, Yale University, New Haven, CT, United States
[6]McGill University, Montreal, QC, Canada
[7]Digitalmill, Freeport, ME, United States
[8]University of Pretoria, Department of Psychology, Pretoria, South Africa

Corresponding Author:
Lynn E Fiellin, MD
play2PREVENT Lab
Yale School of Medicine, Yale University
367 Cedar Street
New Haven, CT, 06510
United States
Phone: 1 203 737 3347
Fax: 1 203 737 3306
Email: lynn.fiellin@yale.edu

Abstract

Background: Human immunodeficiency virus (HIV) disproportionately impacts minority youth. Interventions to decrease HIV sexual risk are needed.

Objective: We hypothesized that an engaging theory-based digital health intervention in the form of an interactive video game would improve sexual health outcomes in adolescents.

Methods: Participants aged 11 to 14 years from 12 community afterschool, school, and summer programs were randomized 1:1 to play up to 16 hours of an experimental video game or control video games over 6 weeks. Assessments were conducted at 6 weeks and at 3, 6, and 12 months. Primary outcome was delay of initiation of vaginal/anal intercourse. Secondary outcomes included sexual health attitudes, knowledge, and intentions. We examined outcomes by gender and age.

Results: A total of 333 participants were randomized to play the intervention (n=166) or control games (n=167): 295 (88.6%) were racial/ethnic minorities, 177 (53.2%) were boys, and the mean age was 12.9 (1.1) years. At 12 months, for the 258 (84.6%) participants with available data, 94.6% (122/129) in the intervention group versus 95.4% (123/129) in the control group delayed initiation of intercourse (relative risk=0.99, 95% CI 0.94-1.05, *P*=.77). Over 12 months, the intervention group demonstrated improved sexual health attitudes overall compared to the control group (least squares means [LS means] difference 0.37, 95% CI 0.01-0.72, *P*=.04). This improvement was observed in boys (LS means difference 0.67, *P*=.008), but not girls (LS means difference 0.06, *P*=.81), and in younger (LS means difference 0.71, *P*=.005), but not older participants (LS means difference 0.03, *P*=.92). The intervention group also demonstrated increased sexual health knowledge overall (LS means difference 1.13, 95% CI 0.64-1.61, *P*<.001), in girls (LS means difference 1.16, *P*=.001), boys (LS means difference 1.10, *P*=.001), younger (LS means difference 1.18, *P*=.001), and older (LS means difference=1.08, *P*=.002) participants. There were no differences in intentions to delay the initiation of intercourse between the two groups (LS means difference 0.10, *P*=.56).

Conclusions: An interactive video game intervention improves sexual health attitudes and knowledge in minority adolescents for at least 12 months.

Trial Registration: Clinicaltrials.gov NCT01666496; https://clinicaltrials.gov/ct2/show/NCT01666496 (Archived by WebCite at http://www.webcitation.org/6syumc9C0).

(J Med Internet Res 2017;19(9):e314) doi:10.2196/jmir.8148

KEYWORDS

adolescent; videogame; intervention; randomized controlled trial; human immunodeficiency virus; risk reduction; primary prevention

Introduction

Background

Human immunodeficiency virus (HIV) disease and sexually transmitted infections (STIs) significantly impact young people, with racial/ethnic minority youth disproportionately affected. Nearly 10,000 youth were diagnosed with HIV in 2014 in the United States, accounting for 22% of new infections [1]. There is a considerable range in both the quantity and quality of education around sexual health, HIV, and acquired immune deficiency syndrome (AIDS) adolescents receive. Only 22 states and the District of Columbia require that public schools teach sex education. Nineteen states require that if sex education is taught, it must be factually accurate [2]. Another 12 states require HIV/AIDS education. Although there are effective HIV and STI prevention programs, challenges in their implementation and fidelity exist. Service providers may lack access to programs [3], selectively implement program components [4], or never implement programs at all [5]. Barriers to implementation include access to adequately trained providers, resource constraints, fidelity, and challenges of adapting an intervention from one population to another [6-8]. To address these issues, digital health interventions have demonstrated efficacy at influencing sexual health [9,10], featuring adaptable content for broader reach, with greater fidelity, at a potentially lower cost [11].

Serious games, defined as video games for a primary purpose beyond pure entertainment [12], offer unique benefits in targeting health promotion and risk prevention [13]. They have efficacy [14] in areas ranging from depression to asthma to cancer [15,16-19]. They also have considerable reach, with 97% of adolescents, including all racial/ethnic groups, playing video games with 50% playing for at least 1 hour or more per day [20]. In a study assessing the impact of digital technologies on schools, more than 333,000 students in grades 6 to 12 reported wanting to use digital games for learning in school and reported games increased their engagement [21]. Active participation through simulated role-playing [22,23] in video games allows individuals to acquire knowledge and practice behavioral skills in a highly accessible, portable, and engaging way [16,24].

Aim of This Study

Our objective was to test the efficacy of a digital health intervention in the form of a theory-driven interactive video game intervention compared to a set of control games on sexual risk behaviors, knowledge, attitudes, and intentions in a population of racial/ethnic minority adolescents. We hypothesized that a highly engaging theory-based video game intervention would have a greater impact on our outcomes of interest than a set of control games.

Methods

Study Design

We conducted a randomized controlled trial in 12 urban community-based settings consisting of seven school-based afterschool programs, four independent afterschool programs, and one summer camp. A description of the development of the PlayForward: Elm City Stories (PlayForward) intervention and the trial design have been published [25,26]. The Yale School of Medicine Human Investigation Committee approved the research.

Participants

Participant eligibility included ages 11 to 14 years, speaks English, able to provide assent and parental/legal guardian consent, and willing to play video games for up to 75 minutes twice weekly for 6 weeks. Although the primary outcome of this study was delay of initiation of sexual intercourse, we enrolled participants who had already initiated intercourse because we did not want potential participants to provide incorrect information regarding their sexual activities to gain or avoid access to the study. These participants (n=6) were not included in these analyses. Participants were provided with an age-appropriate study description and parents/legal guardians were informed that the study focused on promoting healthy behaviors and reducing risk in adolescents. If an adolescent or parent/guardian did not wish to participate, they were not included in the study. All data were collected from 2013 to 2015 and analyzed in 2016.

Randomization

Enrolled participants were randomized in an unmasked fashion to the PlayForward group or to a set of 12 attention-time control video games with all participants playing the games on iPad tablets. A single randomization scheme was generated and written in TrialDB, a customizable Web-based clinical trials database system [27]. After obtaining written youth assent and parental/legal guardian informed consent, eligible participants were assigned to play PlayForward or attention/control games in a 1:1 ratio using a computerized single randomization scheme. Randomization was stratified by gender and age group (11-12 years and 13-14 years), both predictors of the outcomes of interest [28,29]. Notably, gender differences have been identified as being important not only in terms of the trajectory of sexual behaviors in boys and girls during adolescence but also in terms of their response to interventions targeting sexual risk [29]. Randomization was under the control of an investigator who was not involved with eligibility assessment. Study personnel accessed the computerized randomization system, retrieved the assignment, and notified participants of their group assignment.

Study Conditions

Participants played PlayForward or a set of control games on the iPad. PlayForward is a two-dimensional, role-playing adventure video game [25,30-33] developed using a theoretical foundation [32,34-37] about how individual choices made within the social environment of life impact both short-term and long-term goals (Multimedia Appendix 1). These theories primarily focus on the individual's decision making, yet account for the fact that the decisions are made within a certain social context. Delivered as an interactive video game, the intervention allows for the player to see how their individual decisions are influenced by their social surroundings. The game consists of approximately 16 hours of gameplay. The player's goal is to acquire and practice skills to reduce risk behaviors and gain knowledge and healthier attitudes and intentions with the ultimate goal of HIV prevention. The game involves an interactive world where the player creates an Aspirational Avatar [32] and "travels" through life, facing challenges and making decisions in the context of a series of narratives depicting common social situations whose outcomes bring different risks and benefits (Multimedia Appendix 2). The game focuses on sexual health and risk and a range of risky behaviors including substance use, academic dishonesty, and unsafe driving. Five skill-based interactive "mini-games" (Multimedia Appendices 3 and 4) are combined with 12 story-based "challenges" comprising the overarching narratives.

The attention/time control games consisted of 12 video games such as Angry Birds, Dragonbox, and Subway Surfer. They were devoid of content relevant to our study goals and mirrored the number of sessions and length of gameplay in the experimental group [38]. We considered a conventional prevention education control condition; however, there was no "gold standard" for teaching sexual education or HIV prevention in US schools [2], therefore using HIV prevention/education materials as the control was not consistent with current "treatment as usual." We opted for an attention/time control because our primary goal was to determine the efficacy of the video game intervention.

Participants played their assigned game(s) for two sessions per week, approximately 1 hour per session, for 6 weeks on-site at their program [26]. Similarly, all assessments were conducted at the participant's program and included data collected at baseline, 6 weeks (immediately after completion of gameplay), and at 3, 6, and 12 months.

The PlayForward game software records in-game data assessing intervention exposure and fidelity [39]. In-game data documents player's actions, time spent on each action, and overall performance providing a measure of exposure to each intervention component. Periods of inactivity (eg, player is away from the device) can be identified from activity logging data timestamps. Research personnel observed and documented both the PlayForward and the control group participants' duration of gameplay.

Outcome Measures

The primary outcome was delay of initiation of sexual intercourse (defined as initiation of vaginal or anal intercourse) at 12 months post-baseline [40-42]. On June 1, 2015, before any data analyses, the investigators and the Data and Safety Monitoring Board (DSMB) clarified that the primary outcome would be assessed at 12 months postrandomization. Secondary outcomes were sexual attitudes, sexual health knowledge, and sexual intentions. Items assessing the secondary outcomes were compiled from standard instruments and assessed for their internal consistency using Cronbach alpha: delay of the initiation of intercourse [41], sexual attitudes [43] (3 items, maximum score=12, Cronbach alpha=.90), sexual health knowledge [44,45] (15 items, maximum score=15, α=.68), and intentions to delay intercourse [43] (four items, maximum score=16, Cronbach alpha=.84). Although the sexual health knowledge measure (which was piloted before the trial) had a slightly lower level of reliability than is considered acceptable (Cronbach alpha=.7), the items in this measure reflected content that was specifically built into the game, representing very good content validity, which is an important component of a knowledge test in terms of demonstrating that the items within the test are closely related to one another. In addition, although the Cronbach alpha was slightly lower than the .7 value, the result was not borderline, but considerable. Additional assessment data were collected but were not the focus of this paper [26]. Outcomes were assessed at 6 weeks and at 3, 6, and 12 months after randomization. Data were collected face-to-face by research staff and entered into a Web-based database. As described elsewhere [26], we collected data from participants individually; for assessments including sensitive data, participants filled out responses after questions were read to them. Measures were taken to ensure participants felt their responses were being kept confidential.

Statistical Analysis

Power calculations were based on data from published studies [46,47] and systematic reviews [48-50] evaluating youth HIV prevention interventions. These studies suggested a small to moderate effect size of interventions on delaying/preventing initiation of sexual intercourse. National data indicated that approximately 7% of youth reported vaginal or anal intercourse before age 13 years [51] and the final sample size accounted for this figure. A sample size of 330 with 165 in each study condition, was estimated to detect a 15% difference in the proportion of participants achieving the primary outcome of delay of initiation of sexual intercourse (90% PlayForward vs 75% control groups), providing a power of 80% or greater to detect significant (two-tail alpha=.05) differences of this magnitude [52]. This sample size also afforded adequate power (>90%) to detect small to moderate effects on secondary outcome measures. NQuery version 4.0 was used to estimate sample size. This sample size accounted for losses in primary outcome assessment and for enrollment of participants who at baseline had already initiated sexual intercourse.

The primary comparison evaluated the effect of PlayForward compared with the control video games on the delay of initiation of sexual intercourse measured at 12 months postrandomization and constructed as a binary outcome (delayed beyond 12 months/initiated before 12 months). The primary analysis was carried out as an intent-to-treat analysis and adjusted for gender and age (the randomization stratification variables). Sensitivity analyses were carried out with missing responses being assigned both as delay and not delay of initiation of intercourse.

Differences in scores in the secondary outcomes were compared between the two groups at the time points using longitudinal mixed-effects models. Changes in secondary outcome measures since baseline were assessed in repeated measures models (with unstructured covariance), with the assigned baseline values, study group, gender, age, study time point, and study group*time interactions used as covariates. Least squares (LS) means and standard error were plotted for each secondary outcome at each time point. Statistical analyses were done using SAS version 9.4 (SAS Institute Inc, Cary, NC, USA). All protocols were reviewed at intervals by a DSMB.

Results

Description of Study Sample

A total of 333 participants were recruited and enrolled into the study between February 26, 2013 and May 16, 2014; 166 were assigned to the PlayForward intervention and 167 were assigned to the control condition (Figure 1). A total of 166 participants were assigned to PlayForward (162/166, 97.6% initiated gameplay) and 167 to the control games (159/167, 95.2% initiated gameplay). Reasons for not initiating gameplay included participants' inability to participate in the afterschool program due to transportation or medical issues. The PlayForward group played a median of 10.1 hours (interquartile range (IQR) 3.9) over 10 sessions for a median of 60.4 minutes/session (IQR 11.6). The control group played a median of 10.1 hours (IQR 3.9) over 10 sessions for a median of 61.4 minutes/session (IQR 10.9). Eighteen participants withdrew leaving 315 participants in active follow-up. A total of 269 (82.7%) completed the 6-week assessments, 267 (82.9%) completed the 3-month assessments, 253 (78.6%) completed the 6-month assessments, and 258 (81.6%) completed 12-month assessments. No demographic or clinical differences by intervention and control conditions were observed at baseline (Table 1).

Outcomes

The primary outcome was delay of initiation of sexual intercourse. Six participants (PlayForward: n=4; control: n=2) who had engaged in intercourse before baseline were removed from the primary analysis because they had already reached the primary outcome (Table 2).

Overall, the rates of initiating sexual intercourse were low in both groups. There were no differences in rates of delaying initiation of intercourse at each time point in the PlayForward versus the control groups. At 12 months, for the 258 of 304 (84.9%) participants from whom data were available, 122 of 129 (94.6%, 95% CI 89.1%-97.8%) in the PlayForward group versus 123 of 129 (95.4%, 95% CI 90.2%-98.3%) in the control group had delayed the initiation of intercourse (relative risk=0.99, 95% CI 0.94-1.05, *P*=.77). Sensitivity analysis suggested no difference in these proportions (missing=delay, *P*=.90; missing=no delay, *P*=.78).

Over the 12-month follow-up period, the PlayForward group demonstrated an improvement in attitudes about sexual health compared to the control group (LS means difference 0.37, 95% CI 0.01-0.72, *P*=.03) (Figure 2). Treatment*age group and treatment*gender interactions were not significant (*P*=.06 and *P*=.09, respectively). The improvement in attitudes was not seen in girls (LS means difference 0.06, 95% CI –0.46 to 0.58, *P*=.81), but was seen in boys (LS means difference 0.67, 95% CI 0.18-1.16, *P*=.008). In addition, this improvement was only seen in the younger group of participants (age 11-12 years: LS means difference 0.71, 95% CI 0.21-1.20, *P*=.005) and not the older group of participants (age 13-14 years: LS means difference 0.03, 95% CI –0.49 to 0.54, *P*=.92).

Over the 12-month follow-up period, the PlayForward group demonstrated an increase in sexual health knowledge compared to the control group (LS means difference 1.13, 95% CI 0.64-1.61, *P*<.001) (Figure 2). Both girls and boys in the PlayForward group demonstrated an increase in sexual health knowledge compared to the control group (girls: overall LS means difference 1.16, 95% CI 0.46-1.86, *P*=.001; boys: LS means difference 1.10, 95% CI 0.43-1.77, *P*=.001). Both younger (age 11-12 years) and older (age 13-14 years) participants in the PlayForward group demonstrated an increase in sexual health knowledge compared with the control group (younger: LS means difference 1.18, 95% CI 0.50-1.85, *P*=.001; older: LS means difference 1.08, 95% CI 0.39-1.78, *P*=.002).

Over the 12-month follow-up period, there were no differences in intentions to delay the initiation of intercourse between the two groups (LS means difference 0.10, 95% CI –0.23 to 0.43, *P*=.56) (Figure 2). Similarly, there were no differences noted over the 12-month follow-up period by gender or by age group.

Figure 1. Enrollment and follow-up flow diagram for videogame intervention trial for sexual risk reduction. Note: Assessment refers to primary outcome assessment (delay of initiation of sexual intercourse);% is of active participants. Participants who did not initiate game play were still considered active and assessed for study outcomes. A total of six participants (four in PlayForward; two in control) who had initiated sexual intercourse (per primary outcome definition) at baseline were removed from the analysis of primary outcome.

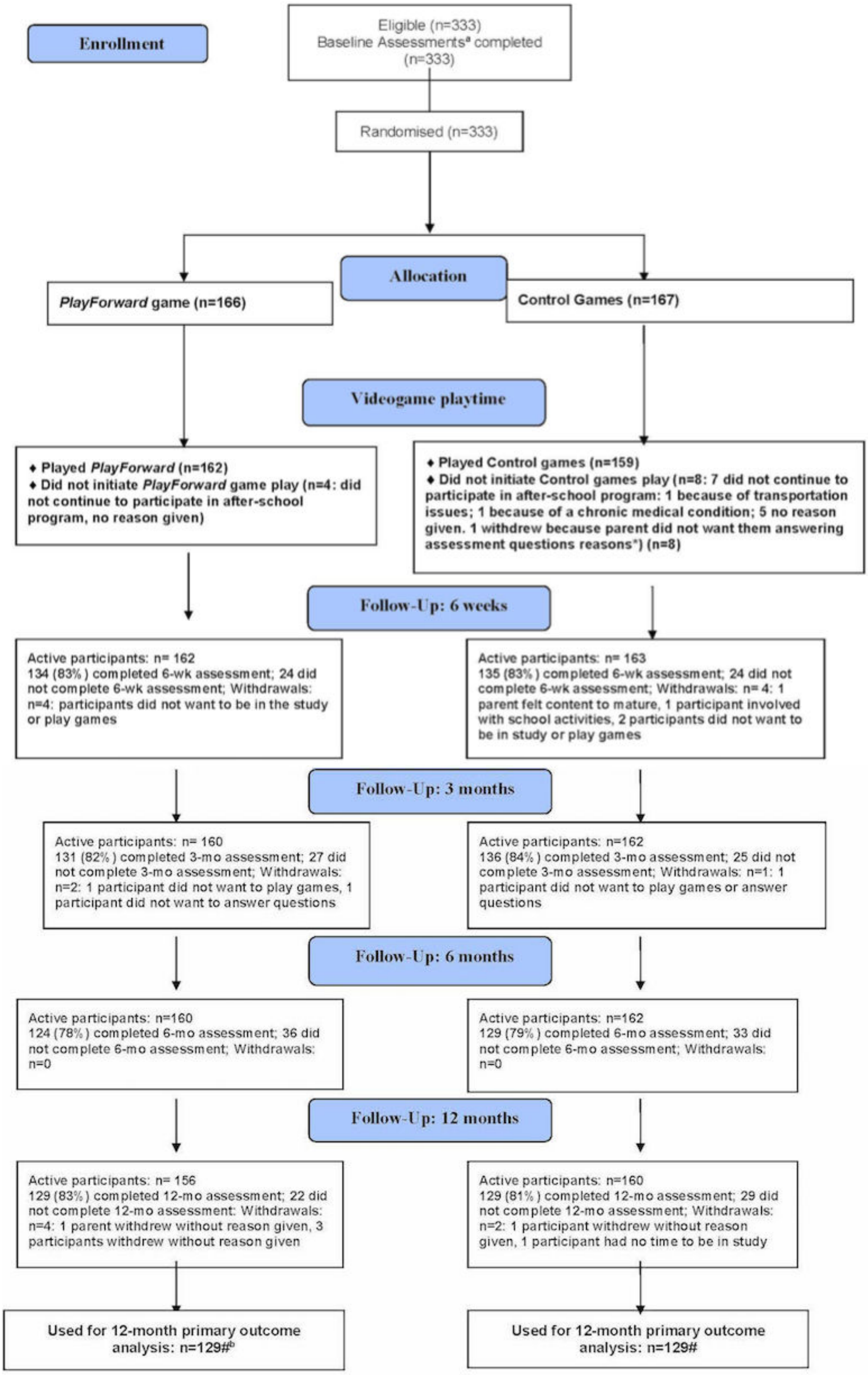


Table 1. Baseline demographic and clinical characteristics of participants.

Characteristics	Control (n=167)	PlayForward (n=166)	Total (N=333)
Gender, n (%)			
Male	89 (53.3)	88 (53.0)	177 (53.2)
Female	78 (46.7)	78 (47.0)	156 (46.8)
Age (years), mean (SD)	12.9 (1.1)	12.9 (1.1)	12.9 (1.1)
Age group (years), n (%)			
11	41 (24.6)	42 (25.3)	83 (24.9)
12	45 (26.9)	45 (27.1)	90 (27.0)
13	45 (26.9)	45 (27.1)	90 (27.0)
14	36 (21.6)	34 (20.5)	70 (21.0)
Race, n (%)			
White	14 (8.6)	17 (10.4)	31 (9.5)
Black	65 (40.1)	70 (42.7)	135 (41.4)
Other	80 (49.4)	76 (46.3)	156 (47.9)
Unknown	3 (1.9)	1 (.6)	4 (1.2)
Ethnicity, n (%)			
Hispanic	87 (55.1)	85 (55.9)	172 (55.5)
Non-Hispanic	71 (44.9)	67 (44.1)	138 (44.5)
Sexual health attitudes score, mean (SD)	10.1 (2.4)	10.1 (2.5)	10.1 (2.5)
Sexual health knowledge score, mean (SD)	6.5 (2.8)	6.1 (2.6)	6.3 (2.7)
Intentions to delay initiation of sex score, mean (SD)	14.9 (2.0)	14.6 (2.2)	14.8 (2.1)

Table 2. Delay of initiation of sexual intercourse by study condition.[a]

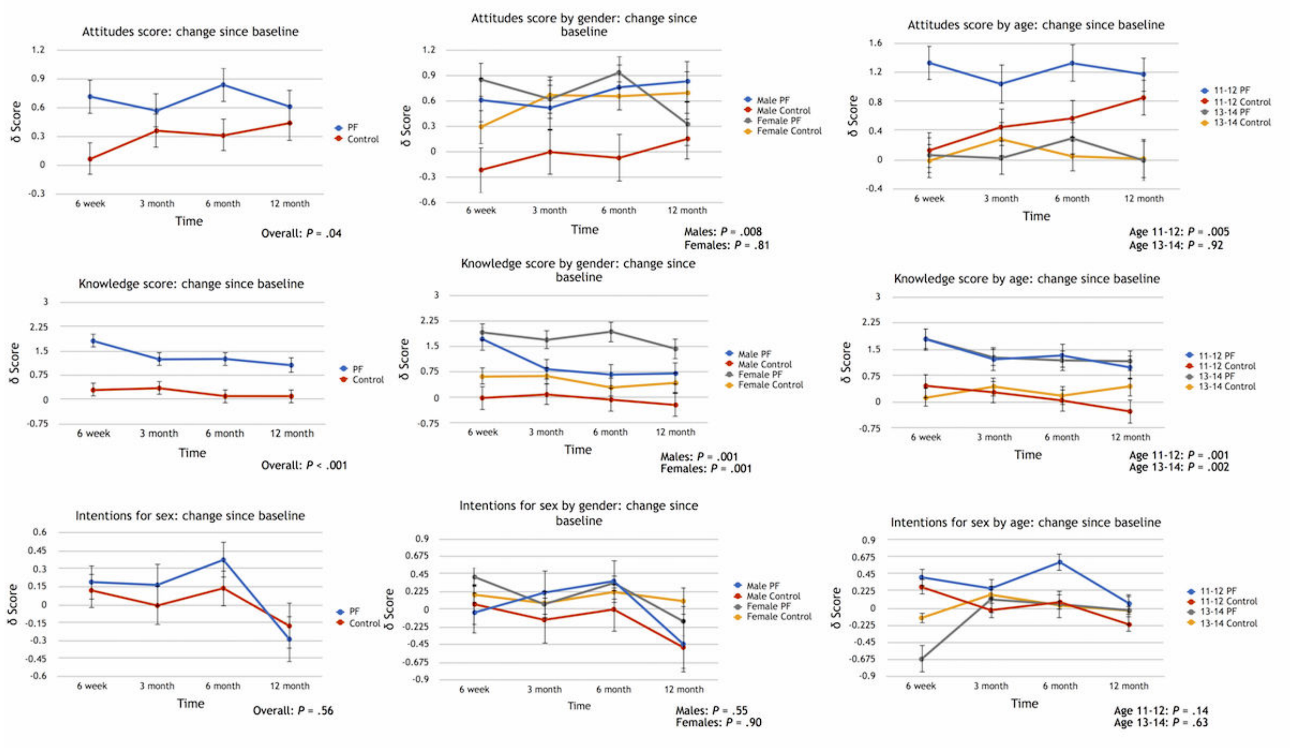
Behavior	Control		PlayForward		Total, n (%)	*P* [c]
	n (%)	95% CI[b]	n (%)	95% CI[b]		
Baseline						
Delay of initiation of sexual intercourse	165 (100)		162 (100)		327 (100)[a]	
6 Weeks						>.99
Delay of initiation of sexual intercourse	132 (97.8)	93.6-99.5	132 (98.5)	94.7-99.8	264 (98.1)	
Initiation of sexual intercourse	3 (1.8)		2 (1.2)		5 (1.5)	
3 Months						.72
Delay of initiation of sexual intercourse	133 (97.8)	93.7-99.54	127 (97.0)	92.4-99.2	260 (97.4)	
Initiation of sexual intercourse	3 (1.8)		4 (2.5)		7 (2.1)	
6 Months						.72
Delay of initiation of sexual intercourse	126 (97.7)	93.4-99.5	120 (96.8)	92.0-99.1	246 (97.2)	
Initiation of sexual intercourse	3 (1.8)		4 (2.5)		7 (2.1)	
12 Months						>.99
Delay of initiation of sexual intercourse	123 (95.4)	90.2-98.3	122 (94.6)	89.1-97.8	245 (95.0)	
Initiation of sexual intercourse	6 (3.6)		7 (4.3)		13 (4.0)	

[a] A total of six participants (control: n=2; PlayForward: n=4), who at baseline were identified (per primary outcome definition) as being engaged in sexual intercourse, were removed from the primary analysis because they had already reached the primary outcome.

[b] Exact 95% confidence intervals are provided for the main outcome (delay of initiation of sexual intercourse).

[c] *P* values are based on a two-sided Fisher exact chi-square test.

Figure 2. Changes in attitudes, knowledge, and intentions by study condition for total group by gender and by age. PF: PlayForward.



Discussion

Principal Findings

In a cohort of community-based racial/ethnic minority adolescents, we found low rates of initiation of sexual intercourse over the 12-month follow-up in the PlayForward and control groups. Those who played PlayForward demonstrated greater improvement in attitudes around sexual health and greater increases in sexual health knowledge over 12 months than the control group.

Interpretation

This study is unique in that it is based in community settings targeting HIV risk in teens using engaging and novel methods that possess the potential for widespread dissemination and impact. To our knowledge, this is the first randomized controlled trial demonstrating the efficacy of a portable sexual health serious video game intervention impacting sexual attitudes and knowledge. The PlayForward game includes many components of programs that have been shown to be effective in reducing sexual risk in adolescents [53,54], such as tailoring programs for the target population, using theory to guide program development, addressing more than just sexual risk, and targeting behaviors most amenable to change. PlayForward was developed with the input of the target audience allowing for its cultural appeal for this population. It incorporates social learning theory and self-efficacy [34,35] and principles from message framing [36] grounded in prospect theory [37]. It addresses myriad risk situations adolescents face, including but not limited to sex and a range of sexual risk behaviors and critical antecedents to those behaviors including attitudes, knowledge, and intentions, with the goal of impacting youth on a number of different levels. It was designed through an iterative approach that focused on our target audience and employed qualitative research methods as well as mixed methods to create and refine the final intervention [55]. Although the theories that serve as the foundation for the game focus on individual reflective processes and the assessments measured these specific processes, the intervention does not disregard automatic processes. Throughout the game, individuals practice their decision-making skills within social contexts. In the narrative sections of the game, the players are challenged to explore the risk around them and learn how that risk influences their individual processes. Where the theory does not specifically target automatic processes, the gameplay does address these processes. Although the game is comprised of up to 16 hours of unique gameplay content, the goal was to expose players to 8 to 12 hours of the intervention given that this amount of time is consistent with existing HIV prevention interventions for adolescents. Athough the median amount of gameplay was 10 hours, many of the players completed the game. An earlier analysis examined the important question as to whether there is a relationship between exposure to different intervention components and study outcomes [39]. This study revealed that it was not the duration of gameplay that seemed to be important in impacting outcomes, but the quality (ie, how well they played) of the participants' performance in specific intervention components of the game.

Our findings are consistent with and complement findings from studies demonstrating the effectiveness of school-based interventions targeting sexual risk reduction in adolescents of similar ages and racial/ethnic backgrounds [42,56,57]. Notably, two of the studies had similar rates of sexual initiation at baseline (4-9%) [42,57] as this study, whereas the third study had considerably higher rates of initiation (23%), indicating it may have been a higher risk population [56]. Similar to this study, one study demonstrated increased positive sexual health attitudes in boys only and increased HIV-related knowledge in boys and girls [57]. This study's intervention delayed sexual initiation and had an impact on other psychosocial determinants in boys, but had a limited effect for girls. The reason for our study's finding of an impact on boys and not girls for sexual health attitudes is not entirely clear. One contributing factor could be that boys, compared to girls, reported feeling more connected to their character in the game (60% vs 39%, P=.01) and this connection could conceivably result in a greater impact on boys' attitudes (unpublished data). Two school-based sexual risk interventions [42,56] demonstrated a delay of initiation of sex in middle-school students, although one did not show an effect in boys or in African-American students [42]. All these studies [42,56,57] required either trained facilitators and/or group discussions as part of the intervention and one study included an additional computer component [42], as compared to this study in which the video game was a stand-alone intervention. In addition, they all followed participants for 24 months, with participants being older at follow-up, allowing for a potentially greater number of events of sexual initiation.

Although there have been long-standing efforts for comprehensive sex education, many challenges for school-based programs remain. Barriers to implementation include competing priorities, lack of parental and administration support, and lack of training [58-61]. In one study, school staff reported confidence in discussing sexuality, but reported varying levels of support for comprehensive sex education from parents (42%), community leaders (53%), and school officials (50%) [60]. In contrast, our qualitative study of key stakeholders implementing PlayForward in real-world settings demonstrated support for the game intervention's potential role in sexual health education in schools [62]. PlayForward offers potential implementation advantages over other school-based interventions because it does not involve intensive training requiring significant human and financial resources, and it provides a level of fidelity in the delivery of content. Importantly, the PlayForward intervention incorporates components of the operational guidance on comprehensive sexuality education promoted by international organizations [63]. These include the provision of scientifically accurate information, a safe learning environment, participatory teaching approaches, strengthening adolescents' skills in communication, decision making, and critical thinking, youth advocacy, and civic engagement in program design and cultural appropriateness, tailored for distinct subpopulations. Technology-based platforms such as the PlayForward video game intervention offer unique advantages because their content and graphics can be rapidly and inexpensively modified and updated, allowing for adaptation of the intervention for different populations and outcomes [64]. Therefore, the PlayForward intervention responds to the call for rigorously evaluated interventions using technology for HIV prevention in adolescents [65]. As the use of digital health interventions is expanding and holds great promise, it is important that they are subject to serious evaluation. Our use of a randomized controlled trial is justified by the ensured stability and engagement of the delivery vehicle of our intervention and that it was delivered with documented fidelity. Given the engagement of our target population in its development and the substantial exposure of our participants to the intervention over time, it has a considerable likelihood of having a clinically meaningful impact [66].

Although we adhered to standards regarding evaluating behavioral interventions [67], some limitations of this study must be considered. A strength of this study is that the intervention was built on a solid foundation of well-established theories, but these theories apply most specifically to behavior change and may have less utility for a prevention intervention where the goal is for the behavior to remain the same. There is considerable precedent in the literature for grounding prevention interventions within a behavior change framework; however, in doing this we may not be accounting for factors that may be relevant to prevention but not behavior change. Notably, the low rate of initiation of sexual intercourse precluded us from determining PlayForward's impact on the primary outcome at 12 months. These low rates are consistent with current national figures [68] with 4% of high school students reporting sexual intercourse before age 13, increasing to 24% by ninth grade (approximately 15 years). Therefore, 12-month follow-up for our cohort may be too short to capture these events. Our outcomes were based on self-report, which can introduce biases in terms of disclosure of sensitive information. As described elsewhere [26], we used data collection methods that optimized disclosure and the accuracy of self-reported data, and ensured privacy and confidentiality in this age group [69,70]. We chose to use face-to-face assessments as we were collecting large amounts of data (15 different assessment instruments) in a young teen cohort and wanted to ensure completion of the assessments. We were cognizant of issues around disclosure given that some of the assessments included sensitive data (sex- and substance use-related data). There are differing views in the literature examining paper versus computer-based assessments of sensitive data [69], but research has found that paper versions elicited higher and more accurate rates of disclosure [71], demonstrated more skipped items, and had no specific advantage to Web-based interviews [70]. Although we studied a population of racial/ethnic minority adolescents who, as a group, are at increased risk for HIV/STIs, they were involved in structured programs and, therefore, the findings may not translate to other populations including those not engaged in afterschool or school-based programs [72].

Despite these limitations and the low event rate of sexual initiation in our cohort, we demonstrated a compelling and persistent impact on attitudes and knowledge. There is evidence for a correlation between attitudes and behavior [73-75], especially in the case of strong attitudes [76]. One study examining survey data from over 1700 ninth graders analyzed the relationship between a number of variables including attitudes and specific outcomes, one being the behavior of abstaining from sexual intercourse [75]. This study found that attitudes had the strongest relationship, not only with intentions (beta=.48) but, more importantly, with behavior (beta=–.34). This finding demonstrated a relationship with students who had high scores indicating more conservative attitudes about having sex being less likely to have sex (P=.01). Furthermore, one meta-analysis concluded that changing attitudes could produce a significant impact specifically on stemming the HIV epidemic

[77]. In addition to digital health interventions such as video games being effective, there remains a critical need for those interventions targeting sexual health to be accessible, adaptable, easily disseminated, and delivered with fidelity. Given the growing focus on the use of digital health (including sexual health [78]) to increase access to and engagement with interventions, this study provides evidence for this approach. According to the 2016 National Education Technology Plan, "technology increasingly is being used to personalize learning and give students more choice over what and how they learn and at what pace, preparing them to organize and direct their own learning for the rest of their lives" and the focus is on "using technology to transform learning experiences with the goal of providing greater equity and accessibility" [79].

Conclusions

Serious video games as digital health interventions offer the unique opportunity to increase the accessibility and reach of theory-driven and tested interventions. The PlayForward intervention demonstrated efficacy in improving sexual attitudes and knowledge in racial/ethnic minority adolescents that persisted up to 12 months. Future research should assess PlayForward's efficacy in populations with higher rates of sexual initiation and its comparative efficacy, effectiveness, and implementation.

Acknowledgments

We acknowledge all authors who contributed significantly to this work. We also acknowledge our game developer partners at Schell Games; our research staff and students, particularly Erica Miller, Orli Florsheim, and Christina Mezes; Gail Slap, MD, MS and Peter Salovey, PhD for their input and support; and Alex Seropian and Noah Falstein for their guidance on the games aspects of this study. We also thank our Data and Safety Monitoring Board members. We offer special thanks to the afterschool and youth programs and their staff and all the adolescents that participated and made a major contribution to this study. The study was supported by grant R01HD062080 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development.

Conflicts of Interest

None declared.

Multimedia Appendix 1

PlayForward homescreen.

[PNG File, 3MB - jmir_v19i9e314_app1.PNG]

Multimedia Appendix 2

A Challenge Stack: 7 Minutes in Heaven.

[PNG File, 3MB - jmir_v19i9e314_app2.PNG]

Multimedia Appendix 3

Know Sense mini-game.

[JPG File, 141KB - jmir_v19i9e314_app3.JPG]

Multimedia Appendix 4

People Sense mini-game.

[PNG File, 1MB - jmir_v19i9e314_app4.PNG]

Multimedia Appendix 5

CONSORT-EHEALTH checklist (v1.6.1).

[PDF File (Adobe PDF File), 768KB - jmir_v19i9e314_app5.pdf]

References

1. Centers for Disease Control and Prevention. 2015 Nov. HIV Surveillance Report, 2014 URL: https://www.cdc.gov/hiv/pdf/library/reports/surveillance/cdc-hiv-surveillance-report-us.pdf [accessed 2016-07-20] [WebCite Cache ID 6qsHNPMh2]

2. National Conference of State Legislatures. State Policies on Sex Education in Schools. 2016 Dec 21 Presented at: National Conference of State Legislatures; May 5, 2015; Chicago, IL URL: http://www.ncsl.org/research/health/state-policies-on-sex-education-in-schools.aspx
3. Collins C, Harshbarger C, Sawyer R, Hamdallah M. The diffusion of effective behavioral interventions project: development, implementation, and lessons learned. AIDS Educ Prev 2006 Aug;18(4 Suppl A):5-20. [doi: 10.1521/aeap.2006.18.supp.5] [Medline: 16987085]
4. Galbraith JS, Stanton B, Boekeloo B, King W, Desmond S, Howard D, et al. Exploring implementation and fidelity of evidence-based behavioral interventions for HIV prevention: lessons learned from the focus on kids diffusion case study. Health Educ Behav 2009 Jun;36(3):532-549 [FREE Full text] [doi: 10.1177/1090198108315366] [Medline: 18445739]
5. Veniegas RC, Kao UH, Rosales R, Arellanes M. HIV prevention technology transfer: challenges and strategies in the real world. Am J Public Health 2009 Apr;99 Suppl 1:S124-S130 [FREE Full text] [doi: 10.2105/AJPH.2007.124263] [Medline: 19218184]
6. Owczarzak J, Dickson-Gomez J. Provider perspectives on evidence-based HIV prevention interventions: barriers and facilitators to implementation. AIDS Patient Care STDS 2011 Mar;25(3):171-179 [FREE Full text] [doi: 10.1089/apc.2010.0322] [Medline: 21323564]
7. Harshbarger C, Simmons G, Coelho H, Sloop K, Collins C. An empirical assessment of implementation, adaptation, and tailoring: the evaluation of CDC's National Diffusion of VOICES/VOCES. AIDS Educ Prev 2006 Aug;18(4 Suppl A):184-197. [doi: 10.1521/aeap.2006.18.supp.184] [Medline: 16987099]
8. Rebchook GM, Kegeles SM, Huebner D. Translating research into practice: the dissemination and initial implementation of an evidence-based HIV prevention program. AIDS Educ Prev 2006 Aug;18(4 Suppl A):119-136. [doi: 10.1521/aeap.2006.18.supp.119] [Medline: 16987094]
9. Bailey JV, Murray E, Rait G, Mercer CH, Morris RW, Peacock R, et al. Interactive computer-based interventions for sexual health promotion. Cochrane Database Syst Rev 2010(9):CD006483. [doi: 10.1002/14651858.CD006483.pub2] [Medline: 20824850]
10. Roberto AJ, Zimmerman RS, Carlyle KE, Abner EL, Cupp PK, Hansen GL. The effects of a computer-based pregnancy, STD, and HIV prevention intervention: a nine-school trial. Health Commun 2007;21(2):115-124. [doi: 10.1080/10410230701306990] [Medline: 17523857]
11. Barak A, Fisher WA. Toward an internet‐driven, theoretically‐based, innovative approach to sex education. J Sex Res 2001 Nov;38(4):324-332. [doi: 10.1080/00224490109552103]
12. Abt C. Serious Games. New York: The Viking Press; 1970.
13. Fiellin L, Hieftje K, Duncan L. Videogames, here for good. Pediatrics 2014 Nov;134(5):849-851 [FREE Full text] [doi: 10.1542/peds.2014-0941] [Medline: 25287452]
14. Hieftje K, Edelman EJ, Camenga DR, Fiellin LE. Electronic media-based health interventions promoting behavior change in youth: a systematic review. JAMA Pediatr 2013 Jun;167(6):574-580 [FREE Full text] [doi: 10.1001/jamapediatrics.2013.1095] [Medline: 23568703]
15. Merry SN, Stasiak K, Shepherd M, Frampton C, Fleming T, Lucassen MFG. The effectiveness of SPARX, a computerised self help intervention for adolescents seeking help for depression: randomised controlled non-inferiority trial. BMJ 2012 Apr 19;344(apr18 3):e2598-e2598. [doi: 10.1136/bmj.e2598]
16. Lieberman D. Interactive video games for health promotion: effects on knowledge, self-efficacy, social support, and health. In: Street RL, Gold WR, Manning T, editors. Health Promotion and Interactive Technology: Theoretical Applications and Future Directions. Mahwah, NJ: Lawrence Erlbaum Associates; 1997:103-120.
17. Lieberman DA. Management of chronic pediatric diseases with interactive health games: theory and research findings. J Ambul Care Manage 2001 Jan;24(1):26-38. [Medline: 11189794]
18. Krishna S, Francisco BD, Balas EA, König P, Graff GR, Madsen RW. Internet-enabled interactive multimedia asthma education program: a randomized trial. Pediatrics 2003 Mar;111(3):503-510. [Medline: 12612228]
19. 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 [FREE Full text] [doi: 10.1542/peds.2007-3134] [Medline: 18676516]
20. Lenhart A, Kahne J, Middaugh E, Macgill A, Evans C, Vitak J. Pew Research Center. 2008 Sep 16. Teens, video games, and civics URL: http://www.pewinternet.org/2008/09/16/teens-video-games-and-civics/ [accessed 2017-08-26] [WebCite Cache ID 6t0gPHQCv]
21. Speak Up Digital Learning 24/7: Understanding Technology-Enhanced Learning in the Lives of Today's Students. Irvine, CA: Project Tomorrow; 2015. URL: http://www.tomorrow.org/speakup/pdfs/SU14StudentReport.pdf [accessed 2017-05-31] [WebCite Cache ID 6qsSpArTl]
22. Shaffer D, Squire K, Halverson R, Gee J. Video games and the future of learning. Phi Delta Kappan 2005;87:105-111 [FREE Full text]
23. Lave J, Wenger E. Situated Learning: Legitimate Peripheral Participation. New York: Cambridge University Press; 1991.

24. Street R, Rimal R. Health promotion and interactive technology: a conceptual foundation. In: Street R, Gold W, Manning T, editors. Health Promotion and Interactive Technology: Theoretical Applications and Future DirectionsI. Mahwah, NJ: Lawrence Erlbaum Associates; 1997:1-18.
25. Hieftje K, Duncan L, Pendergrass T, Sawyer B, Fiellin L. Development of a HIV prevention videogame: lessons learned. Int J Serious Games 2016;3(2):83-90.
26. Fiellin LE, Kyriakides TC, Hieftje KD, Pendergrass TM, Duncan LR, Dziura JD, et al. The design and implementation of a randomized controlled trial of a risk reduction and human immunodeficiency virus prevention videogame intervention in minority adolescents: PlayForward: Elm City Stories. Clin Trials 2016 Aug;13(4):400-408. [doi: 10.1177/1740774516637871] [Medline: 27013483]
27. Nadkarni PM, Brandt C, Frawley S, Sayward FG, Einbinder R, Zelterman D, et al. Managing attribute--value clinical trials data using the ACT/DB client-server database system. J Am Med Inform Assoc 1998;5(2):139-151 [FREE Full text] [Medline: 9524347]
28. Stanton BF, Li X, Ricardo I, Galbraith J, Feigelman S, Kaljee L. A randomized, controlled effectiveness trial of an AIDS prevention program for low-income African-American youths. Arch Pediatr Adolesc Med 1996 Apr;150(4):363-372. [Medline: 8634730]
29. Rotheram-Borus MJ, Song J, Gwadz M, Lee M, Van RR, Koopman C. Reductions in HIV risk among runaway youth. Prev Sci 2003 Sep;4(3):173-187. [Medline: 12940468]
30. Hieftje K, Duncan LR, Fiellin LE. Novel methods to collect meaningful data from adolescents for the development of health interventions. Health Promot Pract 2014 Sep;15(5):714-722 [FREE Full text] [doi: 10.1177/1524839914521211] [Medline: 24519998]
31. Hieftje K, Rosenthal M, Camenga D, Edelman E, Fiellin L. A qualitative study to inform the development of a videogame for adolescent human immunodeficiency virus prevention. Games Health J 2012;1(4):1-5 [FREE Full text] [doi: 10.1089/g4h.2012.0025]
32. Duncan LR, Hieftje KD, Culyba S, Fiellin LE. Game playbooks: tools to guide multidisciplinary teams in developing videogame-based behavior change interventions. Transl Behav Med 2014 Mar;4(1):108-116 [FREE Full text] [doi: 10.1007/s13142-013-0246-8] [Medline: 24653781]
33. Camenga DR, Hieftje KD, Fiellin LE, Edelman EJ, Rosenthal MS, Duncan LR. The use of message framing to promote sexual risk reduction in young adolescents: a pilot exploratory study. Health Educ Res 2014 Apr;29(2):360-366 [FREE Full text] [doi: 10.1093/her/cyt156] [Medline: 24452229]
34. Bandura A. Self-efficacy: toward a unifying theory of behavioral change. Psychol Rev 1977;84(2):191-215. [doi: 10.1037/0033-295X.84.2.191]
35. Bandura A. Self-Efficacy: The Exercise of Control. New York: Freeman; 1997.
36. Rothman AJ, Salovey P. Shaping perceptions to motivate healthy behavior: the role of message framing. Psychol Bull 1997 Jan;121(1):3-19. [Medline: 9000890]
37. Tversky A, Kahneman D. The framing of decisions and the psychology of choice. Science 1981 Jan 30;211(4481):453-458. [Medline: 7455683]
38. Kazdin A. Research Design in Clinical Psychology. New York: Harper & Row; 1980.
39. Montanaro E, Fiellin L, Fakhouri T, Kyriakides T, Duncan L. Using videogame apps to assess gains in adolescents' substance use knowledge: new opportunities for evaluating intervention exposure and content mastery. J Med Internet Res 2015 Oct 28;17(10):e245 [FREE Full text] [doi: 10.2196/jmir.4377] [Medline: 26510775]
40. Centers for Disease Control and Prevention. 2009. 2009 State and Local Youth Risk Behavior Survey URL: http://www.cdc.gov/HealthyYouth/yrbs/pdf/questionnaire/2009HighSchool.pdf [accessed 2008-11-16] [WebCite Cache ID 6qsVJusAA]
41. Centers for Disease Control and Prevention. 2010. 2009 Middle School Youth Risk Behavior Survey URL: http://www.cdc.gov/HealthyYouth/yrbs/pdf/questionnaire/2009MiddleSchool.pdf [accessed 2008-11-16] [WebCite Cache ID 6qsVOOO2g]
42. Tortolero SR, Markham CM, Peskin MF, Shegog R, Addy RC, Escobar-Chaves SL, et al. It's Your Game: Keep It Real: delaying sexual behavior with an effective middle school program. J Adolesc Health 2010 Feb;46(2):169-179 [FREE Full text] [doi: 10.1016/j.jadohealth.2009.06.008] [Medline: 20113923]
43. Basen-Engquist K, Mâsse LC, Coyle K, Kirby D, Parcel GS, Banspach S, et al. Validity of scales measuring the psychosocial determinants of HIV/STD-related risk behavior in adolescents. Health Educ Res 1999 Feb;14(1):25-38. [Medline: 10537945]
44. Kelly JA, St Lawrence JS, Hood HV, Brasfield TL. An objective test of AIDS risk behavior knowledge: scale development, validation, and norms. J Behav Ther Exp Psychiatry 1989 Sep;20(3):227-234. [Medline: 2632590]
45. Jemmott L, Jemmott J, McCaffree K. The National Campaign to Prevent Teen and Unplanned Pregnancy. 2005. Making a Difference!: An abstinence approach to HIV/STD and teen pregnancy prevention URL: http://thenationalcampaign.org/effective-programs/making-difference-abstinence-based-approach-hivstd-and-teen-pregnancy-prevention [accessed 2017-08-26] [WebCite Cache ID 6t0hpHUK4]
46. St Lawrence JS, Brasfield TL, Jefferson KW, Alleyne E, O'Bannon RE, Shirley A. Cognitive-behavioral intervention to reduce African American adolescents' risk for HIV infection. J Consult Clin Psychol 1995 Apr;63(2):221-237. [Medline: 7751483]

47. Clark L, Miller K, Nagy S, Avery J, Roth D, Liddon N, et al. Adult identity mentoring: reducing sexual risk for African-American seventh grade students. J Adolesc Health 2005 Oct;37(4):337. [doi: 10.1016/j.jadohealth.2004.09.024] [Medline: 16182145]
48. Ingram BL, Flannery D, Elkavich A, Rotheram-Borus MJ. Common processes in evidence-based adolescent HIV prevention programs. AIDS Behav 2008 May;12(3):374-383 [FREE Full text] [doi: 10.1007/s10461-008-9369-1] [Medline: 18330687]
49. Johnson BT, Carey MP, Marsh KL, Levin KD, Scott-Sheldon LAJ. Interventions to reduce sexual risk for the human immunodeficiency virus in adolescents, 1985-2000: a research synthesis. Arch Pediatr Adolesc Med 2003 Apr;157(4):381-388. [doi: 10.1001/archpedi.157.4.381] [Medline: 12695235]
50. Pedlow C, Carey M. HIV sexual risk-reduction interventions for youth: A review and methodological critique of randomized controlled trials. Behav Modif 2003;27:135-190 [FREE Full text]
51. Eaton DK, Kann L, Kinchen S, Shanklin S, Ross J, Hawkins J, et al. Youth risk behavior surveillance--United States, 2007. MMWR Surveill Summ 2008 Jun 06;57(4):1-131 [FREE Full text] [Medline: 18528314]
52. Cohen J. Statistical Power Analysis for Behavioral Sciences. Hillsdale, NJ: Lawrence Erlbaum Associates; 1988.
53. Goesling B, Colman S, Trenholm C, Terzian M, Moore K. Programs to reduce teen pregnancy, sexually transmitted infections, and associated sexual risk behaviors: a systematic review. J Adolesc Health 2014 May;54(5):499-507. [doi: 10.1016/j.jadohealth.2013.12.004] [Medline: 24525227]
54. Sales JM, Milhausen RR, Diclemente RJ. A decade in review: building on the experiences of past adolescent STI/HIV interventions to optimise future prevention efforts. Sex Transm Infect 2006 Dec;82(6):431-436 [FREE Full text] [doi: 10.1136/sti.2005.018002] [Medline: 17151029]
55. Yardley L, Spring BJ, Riper H, Morrison LG, Crane DH, Curtis K, et al. Understanding and promoting effective engagement with digital behavior change interventions. Am J Prev Med 2016 Nov;51(5):833-842. [doi: 10.1016/j.amepre.2016.06.015] [Medline: 27745683]
56. Jemmott JB, Jemmott LS, Fong GT. Efficacy of a theory-based abstinence-only intervention over 24 months: a randomized controlled trial with young adolescents. Arch Pediatr Adolesc Med 2010 Feb;164(2):152-159 [FREE Full text] [doi: 10.1001/archpediatrics.2009.267] [Medline: 20124144]
57. Coyle KK, Kirby DB, Marín BV, Gómez CA, Gregorich SE. Draw the line/respect the line: a randomized trial of a middle school intervention to reduce sexual risk behaviors. Am J Public Health 2004 May;94(5):843-851. [Medline: 15117710]
58. Fagen MC, Stacks JS, Hutter E, Syster L. Promoting implementation of a school district sexual health education policy through an academic-community partnership. Public Health Rep 2010;125(2):352-358 [FREE Full text] [doi: 10.1177/003335491012500227] [Medline: 20297764]
59. Alldred P, David M, Smith P. Teachers' views of teaching sex education: pedagogy and models of delivery. J Educ Enquiry 2003;4(1):80-96.
60. Peskin M, Hernandez B, Markham C, Johnson K, Tyrell S, Addy R, et al. Sexual health education from the perspective of school staff: implications for adoption and implementation of effective programs in middle school. J Applied Res Children 2011;2(2):9-38.
61. Donovan P. School-based sexuality education: the issues and challenges. Fam Plann Perspect 1998;30(4):188-193. [doi: 10.1363/3018898]
62. Pendergrass TM, Hieftje K, Crusto CA, Montanaro E, Fiellin LE. If we build it, will they come? A qualitative study of key stakeholder opinions on the implementation of a videogame intervention for risk reduction in adolescents. Games Health J 2016 Aug;5(4):279-285 [FREE Full text] [doi: 10.1089/g4h.2015.0092] [Medline: 27336205]
63. UNFPA Operational Guidance for Comprehensive Sexuality Education: A Focus on Human Rights and Gender. New York: United Nations Population Fund; 2014. URL: http://www.unfpa.org/sites/default/files/pub-pdf/UNFPA_OperationalGuidance_WEB3.pdf [accessed 2016-05-02] [WebCite Cache ID 6qsYldFKC]
64. Oldenburg B, Taylor CB, O'Neil A, Cocker F, Cameron LD. Using new technologies to improve the prevention and management of chronic conditions in populations. Annu Rev Public Health 2015 Mar 18;36:483-505. [doi: 10.1146/annurev-publhealth-031914-122848] [Medline: 25581147]
65. Pettifor A, Bekker L, Hosek S, DiClemente R, Rosenberg M, Bull SS, et al. Preventing HIV among young people: research priorities for the future. J Acquir Immune Defic Syndr 2013 Jul;63 Suppl 2:S155-S160 [FREE Full text] [doi: 10.1097/QAI.0b013e31829871fb] [Medline: 23764629]
66. Murray E, Hekler EB, Andersson G, Collins LM, Doherty A, Hollis C, et al. Evaluating digital health interventions: key questions and approaches. Am J Prev Med 2016 Nov;51(5):843-851. [doi: 10.1016/j.amepre.2016.06.008] [Medline: 27745684]
67. Boutron I, Moher D, Altman DG, Schulz KF, Ravaud P. Extending the CONSORT statement to randomized trials of nonpharmacologic treatment: explanation and elaboration. Ann Intern Med 2008 Feb 19;148(4):295-309. [Medline: 18283207]
68. Centers for Disease Control and Prevention. 2016. Youth Risk Behavior Surveillance -- United States, 2015 URL: https://www.cdc.gov/mmwr/volumes/65/ss/ss6506a1.htm [accessed 2016-11-01] [WebCite Cache ID 6qsXI5HbE]

69. Flanagan SM, Greenfield S, Coad J, Neilson S. An exploration of the data collection methods utilised with children, teenagers and young people (CTYPs). BMC Res Notes 2015 Mar 01;8:61 [FREE Full text] [doi: 10.1186/s13104-015-1018-y] [Medline: 25888787]
70. Wyrick DL, Bond L. Reducing sensitive survey response bias in research on adolescents: a comparison of web-based and paper-and-pencil administration. Am J Health Promot 2011;25(5):349-352. [doi: 10.4278/ajhp.080611-ARB-90] [Medline: 21534838]
71. Beebe T, Harrison P, Park E, Evans J. The effects of data collection mode and disclosure on adolescent reporting of health behaviour. Soc Sci Comput Rev 2006;24(4):476-488 [FREE Full text]
72. Afterschool Alliance. America After 3PM: Afterschool Programs in Demand. 2014. URL: http://www.afterschoolalliance.org/documents/AA3PM-2014/AA3PM_Key_Findings.pdf [accessed 2017-05-31] [WebCite Cache ID 6qsWvNtUX]
73. Albarracin D, Durantini M, Earl A. Empirical and theoretical conclusions of an analysis of outcomes of HIV-prevention interventions. Curr Dir Psychol Sci 2006;15(2):73-78 [FREE Full text]
74. Ajzen I, Fishbein M. The influence of attitudes on behavior. In: Albarracin D, Johnson BT, Zanna MP, editors. The Handbook of Attitudes. Mahwah, NJ: Lawrence Erlbaum Associates; 2005:173-221.
75. Basen-Engquist K, Parcel G. Attitudes, norms, and self-efficacy: a model of adolescents' HIV-related sexual risk behavior. Health Educ Q 1992;19(2):263-277. [Medline: 1618632]
76. Holland RW, Verplanken B, Van Knippenberg A. On the nature of attitude-behavior relations: the strong guide, the weak follow. Eur J Soc Psychol 2002 Nov;32(6):869-876. [doi: 10.1002/ejsp.135]
77. Albarracín D, Johnson BT, Fishbein M, Muellerleile PA. Theories of reasoned action and planned behavior as models of condom use: a meta-analysis. Psychol Bull 2001 Jan;127(1):142-161 [FREE Full text] [Medline: 11271752]
78. Allison S, Bauermeister JA, Bull S, Lightfoot M, Mustanski B, Shegog R, et al. The intersection of youth, technology, and new media with sexual health: moving the research agenda forward. J Adolesc Health 2012 Sep;51(3):207-212 [FREE Full text] [doi: 10.1016/j.jadohealth.2012.06.012] [Medline: 22921129]
79. US Department of Education, Office of Educational Technology. 2016 National Education Technology Plan. 2016. Future ready learning: reimagining the role of technology in education URL: https://tech.ed.gov/files/2015/12/NETP16.pdf [accessed 2017-05-31] [WebCite Cache ID 6qsVnbAUj]

Abbreviations

AIDS: acquired immune deficiency syndrome
DSMB: data and safety monitoring board
HIV: human immunodeficiency virus
LS: least squares
STI: sexually transmitted infection

Edited by G Eysenbach; submitted 04.06.17; peer-reviewed by K Curtis, A Fitzgerald; comments to author 06.07.17; revised version received 09.08.17; accepted 14.08.17; published 18.09.17.

Please cite as:
Fiellin LE, Hieftje KD, Pendergrass TM, Kyriakides TC, Duncan LR, Dziura JD, Sawyer BG, Mayes L, Crusto CA, Forsyth BWC, Fiellin DA
Video Game Intervention for Sexual Risk Reduction in Minority Adolescents: Randomized Controlled Trial
J Med Internet Res 2017;19(9):e314
URL: http://www.jmir.org/2017/9/e314/
doi:10.2196/jmir.8148
PMID:28923788

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org