Journal of Medical Internet Research

Journal Impact Factor (JIF) (2023): 5.8
Volume 18 (2016), Issue 5 ISSN 1438-8871 Editor in Chief: Gunther Eysenbach, MD, MPH

Contents

Corrigenda and Addenda

Viewpoint

The Emergence of Personalized Health Technology

Luke Nelson Allen[1], BSc (Hon), MPH, MBChB, PGCE; Gillian Pepall Christie[2], MA, MPhil

[1]Harvard T.H. Chan School of Public Health, Department of Global Health and Population, Boston, MA, United States
[2]Vitality Group, New York, NY, United States

Corresponding Author:
Luke Nelson Allen, BSc (Hon), MPH, MBChB, PGCE
Harvard T.H. Chan School of Public Health
Department of Global Health and Population
677 Huntington Av
Boston, MA, Ma 02115
United States
Phone: 1 1865289471
Fax: 1 1865289472
Email: luke.allen@mail.harvard.edu

Abstract

Personalized health technology is a noisy new entrant to the health space, yet to make a significant impact on population health but seemingly teeming with potential. Devices including wearable fitness trackers and healthy-living apps are designed to help users quantify and improve their health behaviors. Although the ethical issues surrounding data privacy have received much attention, little is being said about the impact on socioeconomic health inequalities. Populations who stand to benefit the most from these technologies are unable to afford, access, or use them. This paper outlines the negative impact that these technologies will have on inequalities unless their user base can be radically extended to include vulnerable populations. Frugal innovation and public–private partnership are discussed as the major means for reaching this end.

(J Med Internet Res 2016;18(5):e99) doi:10.2196/jmir.5357

KEYWORDS

personalized health technology; population health; frugal innovation; ethics; socioeconomic factors, inequalities; technology, health

Introduction

Several larger technology giants and smaller upstarts are creating personalized health technologies. Sensors, smartwatches, and mobile health apps are strapped to wrists and placed in pockets to monitor and help to modify health behaviors [1]. High-profile devices include the Fitbit, Jawbone, Microsoft Band, and Apple smartwatch. While there is no widely agreed-upon definition, personalized health technology generally refers to wearable devices that monitor health-related activity and provide feedback at the individual level, usually through a corresponding app or minidisplay on the device. Current products track lifestyle information such as steps walked, hours slept, and calories consumed. Terabytes of data are analyzed to deliver instantaneous and predictive insights to users. These technologies contribute to the self-quantification movement and to the consumerization of health.

PHTs that empower consumers to quantify health behaviors could advance health for all populations. The modifiable risk factors that are measured by these devices—including physical activity and diet—are major drivers of noncommunicable diseases. These conditions, including cardiovascular and lung diseases, type 2 diabetes, and various cancers, are the leading causes of death and disability worldwide [2]. They account for an estimated 60% of all deaths and will cost the global economy US $30 trillion by 2025 [2,3].

Scientific evidence demonstrating the impact of personalized health technologies on health is still emerging. This is partly because rapid technological innovation is being driven by firms responding to a market for these products, leaving academics and clinicians to play catch-up with health impact evaluations. Early research suggests that the technologies can facilitate changes in behaviors and reductions in disease risks, and that the health impact is magnified when coupled with broader engagement strategies [4-6]. Although noncommunicable diseases disproportionately afflict disadvantaged groups [2,7,8], the uptake of personalized health technologies has been limited to the educated, healthy, and wealthy [9-11]. Devices remain largely unaffordable and inaccessible to lower-income populations, and many lack the technological skills required to

modify their health behaviors using information delivered by the technologies [12-14]. Until access is expanded to nontraditional users, personalized health technologies will continue to widen socioeconomic health inequities for vulnerable populations worldwide.

The Problem: Exacerbating Inequity

Inequities are often exacerbated in the shorter term when innovative technologies enter the marketplace. Companies place higher prices on new technologies to recover their original investment in research and development. Over time, the power of economies of scale and competitive forces push prices downward: technologies eventually become affordable to the masses. As relatively new products, personalized health technologies have served to widen inequities because only affluent early adopters can afford their higher prices, while marginalized populations remain excluded [10,15].

Personalized health technologies further exacerbate inequities in the shorter term because early adopters are motivated and health conscious [11]. Existing users tend to be highly educated and possess the necessary technological skills to operate the devices. They also have the linguistic and numeric capabilities to process information in order to change behaviors. While high-income earners possess these skills, vulnerable populations—older adults, racial and ethnic minorities, poorly educated individuals, and low-income earners—commonly lack them [16]. Without adequate technological or health literacy skills, marginalized populations cannot actively engage with personalized health technologies. These factors may widen socioeconomic health inequities further in the short term [15].

How can gaps in the affordability and accessibility of personalized health technologies be closed in the longer term? According to Tudor Hart's inverse care law, health products and services are always used most by those who need them least [17]. As devices become more efficacious, advancements in health will continue to disproportionately benefit the privileged. Personalized health technology will not realize its public health potential in reducing the global burden of noncommunicable diseases unless challenges associated with the affordability and accessibility of personalized health technologies are proactively mitigated.

The Solution: Frugal Innovation

Creative strategies are required to advance health for individuals occupying lower rungs of the socioeconomic ladder. Innovating for the bottom of the pyramid—the 3 billion people living on less than US $2.50 a day—is not new [18]. In 2002, the renowned management professor CK Prahalad proposed the development of products and services for the bottom of the pyramid. Prahalad realized that fortunes were being left on the table as companies neglected to target the largest but poorest socioeconomic population [19]. Since Prahalad's writings, innovating with a frugal innovation mindset has emerged to target these previously marginalized consumers.

Frugal innovations are high-quality products created with limited resources [20]. Innovating with a frugal innovation mindset entails reducing the cost and complexity of products by removing nonessential features to create "good enough" products. Materials are repeatedly recycled to self-sustain the company and the environment, while diverse external partners such as universities and venture capitalists are often brought together to maximize efficiencies. The needs and requirements of end users in bottom-of-the-pyramid markets are central to the development process.

This relatively new model has received widespread support from socially oriented enterprises and influential corporate leaders alike. Unilever's Chief Executive Officer, Paul Polman, wrote in the foreword of Radjou and colleague's book *Frugal Innovation* that the "frugal ingenuity of developing nations with the advanced [research and development] capabilities of advanced economies [can enable] companies to create high-quality products and services that are affordable, sustainable, and benefit humanity as a whole" [20]. Indra Nooyi, Chairperson and Chief Executive Officer of PepsiCo, further contends that "frugal innovation is one of the most critical emerging models of value creation for both businesses and the customers they serve" [20].

Numerous companies have created products and services for health using this approach [21,22]. General Electric has developed an electrocardiograph machine that costs US $800 as opposed to US $2000 and has reduced the cost of an electrocardiographic test to US $1 per person. Tata has established the Tata Swach to purify water without running water or electricity for US $20. The innovative Jaipur foot is a prosthetic that costs less than US $45 [23]. A majority of these innovations are widely used in developing as well as in developed countries. They also enable developing countries to leapfrog their developed country counterparts to provide cost-effective innovations at scale.

Despite the emergence of frugal innovations, challenges arise that could hinder their broader uptake and use. Predicting what consumers need and desire is relatively easier than actually engaging them, particularly when the target group is a marginalized population. In addition, balancing financial and social returns to sell at scale while continuing to generate profits is a ubiquitous issue for companies engaged in frugal innovation.

It is also important that emerging systems, tools, and personalized health technology devices be subject to rigorous technology assessments [24]. Established qualitative and quantitative tools [25] can be used to evaluate performance against a range of key performance indicators that extend beyond health metrics to include reliability, integration with other devices, cost, and data security [24].

Frugal Innovation and Personalized Health Technology

A frugal innovation mindset can be applied to personalized health technologies to minimize socioeconomic health inequities. Companies can engage end users in bottom-of-the-pyramid markets to design, develop, and test the effectiveness of personalized health technologies. There are several ongoing pilot projects that use personalized health technologies to

improve health outcomes in low-income populations [26-28]. Unfortunately, the majority of early personalized health technology initiatives in low- and middle-income countries tended to lack careful targeting, sustainable funding, robust evaluation, and the ability to work at scale [29-31]. In addition to these projects, more affordable personalized health technologies and programs are emerging to benefit bottom-of-the-pyramid markets, including the Xiaomi Mi Band and United Nations Children's Fund (UNICEF) Kid Power Bands. In a recent study, the Mi Band—sold in Asia and priced at less than US $20—outperformed more expensive competitors in accurately monitoring vital signs and steps walked [32]. Another innovative example of frugal personalized health technology that benefits low-income groups in both high- and low-income countries is the UNICEF Kid Power Program. UNICEF and the US retailer Target market the discounted Kid Power Band (a wrist-worn activity-monitoring device) at children in high-income countries with obesity problems. By meeting physical activity goals, wearers earn points, recorded on a parallel mobile app. Corporate sponsors translate these points into funding for childhood nutrition programs in developing countries. To date, over 50,000 children have engaged with the initiative, raising enough money to provide over a quarter of a million UNICEF therapeutic food packets for malnourished children [33].

As evidence for the effectiveness of these technologies mounts, governments with legal and constitutional obligations to promote the health of their citizens can subsidize the creation of personalized health technologies and provide them through state-funded programs. These measures will help to mitigate the inequities that this disruptive technology is exacerbating.

Conclusion: A Call for Collective Action

Technologies are tools—they are designed by people and for people. While personalized health technologies are pregnant with potential, the extent to which they affect the health of all populations remains limited. With the current user base, any health benefits derived from using personalized health technologies or future iterations will disproportionately accrue to the affluent unless strategies are adopted to widen access among disadvantaged groups. In time, new business models are likely to emerge that reduce costs, increase affordability, and expand access. Until then, the combination of frugal innovation with public and private sector action can leverage personalized health technologies to advance global health responsibly, sustainably, and equitably.

Acknowledgments

We are grateful to Dr Derek Yach, Chief Health Officer at the Vitality Group, for his insightful guidance and comments on earlier versions of this manuscript. We thank Sarah Kunkle for her support in the preparation of this paper.

This paper received no specific funding.

Conflicts of Interest

None declared.

References

1. Lyons EJ, Lewis ZH, Mayrsohn BG, Rowland JL. Behavior change techniques implemented in electronic lifestyle activity monitors: a systematic content analysis. J Med Internet Res 2014 Aug;16(8):e192 [FREE Full text] [doi: 10.2196/jmir.3469] [Medline: 25131661]
2. World Health Organization. Noncommunicable diseases. Geneva, Switzerland: World Health Organization; 2015 Jan. URL: http://www.who.int/mediacentre/factsheets/fs355/en/ [accessed 2015-05-03] [WebCite Cache ID 6d8UOOTir]
3. Bloom DE, Cafiero ET, Jané-Llopis E, Abrahams-Gessel S, Bloom LR, Fathima S, Feigl AB, Gaziano T, Mowafi M, Pandya A, Prettner K, Rosenberg L, Seligman B, Stein AZ, Weinstein C. The Global Economic Burden of Non-communicable Diseases. Geneva: World Economic Forum; 2011 Sep. URL: http://www3.weforum.org/docs/WEF_Harvard_HE_GlobalEconomicBurdenNonCommunicableDiseases_2011.pdf [accessed 2016-04-22] [WebCite Cache ID 6gxcETMst]
4. Hsu WC, Lau KH, Huang R, Ghiloni S, Le H, Gilroy S, et al. Utilization of a cloud-based diabetes management program for insulin initiation and titration enables collaborative decision making between healthcare providers and patients. Diabetes Technol Ther 2016 Feb;18(2):59-67 [FREE Full text] [doi: 10.1089/dia.2015.0160] [Medline: 26645932]
5. Rabbi M, Pfammatter A, Zhang M, Spring B, Choudhury T. Automated personalized feedback for physical activity and dietary behavior change with mobile phones: a randomized controlled trial on adults. JMIR Mhealth Uhealth 2015 May;3(2):e42 [FREE Full text] [doi: 10.2196/mhealth.4160] [Medline: 25977197]
6. Bravata DM, Smith-Spangler C, Sundaram V, Gienger AL, Lin N, Lewis R, et al. Using pedometers to increase physical activity and improve health: a systematic review. JAMA 2007 Nov 21;298(19):2296-2304. [doi: 10.1001/jama.298.19.2296] [Medline: 18029834]
7. World Health Organization. Global Status Report on Alcohol and Health 2014. Geneva, Switzerland: WHO; May 12, 2014:1-392.

8. Mackenbach JP, Stirbu I, Roskam AR, Schaap MM, Menvielle G, Leinsalu M, European Union Working Group on Socioeconomic Inequalities in Health. Socioeconomic inequalities in health in 22 European countries. N Engl J Med 2008 Jun 5;358(23):2468-2481. [doi: 10.1056/NEJMsa0707519] [Medline: 18525043]
9. Patel MS, Asch DA, Volpp KG. Wearable devices as facilitators, not drivers, of health behavior change. JAMA 2015 Feb 3;313(5):459-460. [doi: 10.1001/jama.2014.14781] [Medline: 25569175]
10. Fox S, Duggan M. Tracking for health. Washington, DC: Pew Internet & American Life Project; 2013 Jan 28. URL: http://www.pewinternet.org/files/old-media//Files/Reports/2013/PIP_TrackingforHealth%20with%20appendix.pdf [accessed 2016-02-11] [WebCite Cache ID 6fDMpBEIs]
11. Epps SR, Doty CA, Knott J, Vokshi A. Fitness Wearables — Many Products, Few Customers. Cambridge, MA: Forrester Reserach Inc; 2013 Jan 28. URL: https://www.forrester.com/Fitness+Wearables+Many+Products+Few+Customers/fulltext/-/E-res89941 [accessed 2016-02-05] [WebCite Cache ID 6f4KBG0xx]
12. Ericsson Mobility Report: On the Pulse of the Networked Society. Stockholm, Sweden: Ericsson; 2014 Jun. URL: http://www.ericsson.com/res/docs/2014/ericsson-mobility-report-june-2014.pdf [accessed 2016-02-11] [WebCite Cache ID 6fDN3vJEn]
13. Rhodes H. Accessing and using data from wearable fitness devices. J AHIMA 2014 Sep;85(9):48-50. [Medline: 25286622]
14. Nelson LA, Mulvaney SA, Gebretsadik T, Ho Y, Johnson KB, Osborn CY. Disparities in the use of a mHealth medication adherence promotion intervention for low-income adults with type 2 diabetes. J Am Med Inform Assoc 2016 Jan;23(1):12-18. [doi: 10.1093/jamia/ocv082] [Medline: 26186935]
15. McAuley A. Digital health interventions: widening access or widening inequalities? Public Health 2014 Dec;128(12):1118-1120. [doi: 10.1016/j.puhe.2014.10.008] [Medline: 25458115]
16. Quick Guide to Health Literacy. Washington, DC: U.S. Department of Health and Human Services URL: http://health.gov/communication/literacy/quickguide/factsbasic.htm [accessed 2015-05-03] [WebCite Cache ID 6d8TodiGh]
17. Hart JT. The inverse care law. The Lancet 1971 Feb;297(7696):405-412. [doi: 10.1016/S0140-6736(71)92410-X]
18. Shah A. Global Issues: Poverty Facts and Stats.: Anup Shah; 2013 Jan 07. URL: http://www.globalissues.org/article/26/poverty-facts-and-stats [accessed 2015-11-18] [WebCite Cache ID 6d8TXVlk9]
19. Prahalad CK, Hart SL. The fortune at the bottom of the pyramid. Strategy+Business 2002 Jan 10;First Quarter(26):1-26 [FREE Full text]
20. Radjou N, Prabhu J. Frugal Innovation: How to do More With Less. New York, NY: Public Affairss; 2015.
21. O'Hara NN. Is safe surgery possible when resources are scarce? BMJ Qual Saf 2015 Jul;24(7):432-434. [doi: 10.1136/bmjqs-2015-004377] [Medline: 25998062]
22. Tiwari R, Herstatt C. Assessing India's lead market potential for cost-effective innovations. J Indian Bus Res 2012 Jun;4(2):97-115. [doi: 10.1108/17554191211228029]
23. Bound K, Thornton I. Our Frugal Future: Lessons from India's Innovation System. London, UK: Nesta; 2012 Jul. URL: http://www.nesta.org.uk/sites/default/files/our_frugal_future.pdf [accessed 2016-04-22] [WebCite Cache ID 6gxeRlcly]
24. Martinez-Millana A, Fico G, Fernández-Llatas C, Traver V. Performance assessment of a closed-loop system for diabetes management. Med Biol Eng Comput 2015 Dec;53(12):1295-1303. [doi: 10.1007/s11517-015-1245-3] [Medline: 25667016]
25. van der Weegen S, Verwey R, Spreeuwenberg M, Tange H, van der Weijden T, de Witte L. The development of a mobile monitoring and feedback tool to stimulate physical activity of people with a chronic disease in primary care: a user-centered design. JMIR Mhealth Uhealth 2013 Jul;1(2):e8 [FREE Full text] [doi: 10.2196/mhealth.2526] [Medline: 25099556]
26. Baum S. MedCity News: Omada Health Raises $48M With Plans to Extend Program to Medicaid Patients. New York, NY: Breaking Media Inc; 2015 Sep 16. URL: http://medcitynews.com/2015/09/omada-health-raises-48m-with-plans-to-provide-program-for-medicaid-patients/ [accessed 2016-02-11] [WebCite Cache ID 6fDO544X8]
27. LaMantia J. Crain's New York Business: Verizon Teams Up With Digital Health Company to Give Broadband to Low-Income Patients. New York, NY: Crain Communications Inc; 2015 Jul 23. URL: http://www.crainsnewyork.com/article/20150723/TECHNOLOGY/150729918/verizon-teams-up-with-digital-health-company-to-give-broadband-to-low-income-patients [accessed 2016-01-28] [WebCite Cache ID 6espUr6Np]
28. Winners of the Super-Utilizer Health Innovation Challenge: AdhereTech, A-CHESS, and Wellth. Hamilton, NJ: Center for Health Care Strategies, Inc; 2015 Aug. URL: http://www.chcs.org/highlight/winners-super-utilizer-health-innovation-challenge-adheretech-chess-wellth/ [accessed 2016-01-28] [WebCite Cache ID 6espXMidu]
29. Lemaire J. Scaling Up Mobile Health: Elements Necessary for the Successful Scale Up of mHealth in Developing Countries. Geneva: Advanced Development for Africa; 2011 Dec. URL: https://www.k4health.org/sites/default/files/ADA_mHealth%20White%20Paper.pdf [accessed 2016-04-22] [WebCite Cache ID 6gxdzHmkN]
30. Blaya JA, Fraser HS, Holt B. E-health technologies show promise in developing countries. Health Aff (Millwood) 2010 Feb;29(2):244-251 [FREE Full text] [doi: 10.1377/hlthaff.2009.0894] [Medline: 20348068]
31. Lewis T, Synowiec C, Lagomarsino G, Schweitzer J. E-health in low- and middle-income countries: findings from the Center for Health Market Innovations. Bull World Health Organ 2012 May 1;90(5):332-340 [FREE Full text] [doi: 10.2471/BLT.11.099820] [Medline: 22589566]

32. El-Amrawy F, Nounou M. Are currently available wearable devices for activity tracking and heart rate monitoring accurate, precise, and medically beneficial? Healthc Inform Res 2015 Oct 20;21(4):315-320 [FREE Full text]
33. UNICEF Kid Power: Get Active. Save Lives. New York, NY: U.S. Fund for UNICEF; 2015. URL: http://unicefkidpower.org/ [accessed 2016-02-05] [WebCite Cache ID 6f4FrFxML]

Abbreviations

UNICEF: United Nations Children's Fund

Edited by G Eysenbach; submitted 19.11.15; peer-reviewed by J Williams, A Ottaway, A Martinez-Millana; comments to author 24.12.15; revised version received 10.02.16; accepted 10.02.16; published 10.05.16.

Please cite as:
Allen LN, Christie GP
The Emergence of Personalized Health Technology
J Med Internet Res 2016;18(5):e99
URL: http://www.jmir.org/2016/5/e99/
doi:10.2196/jmir.5357
PMID:27165944

Viewpoint

Medical Tourism and Telemedicine: A New Frontier of an Old Business

Yan Alicia Hong[1,2], PhD

[1]School of Public Health, Xiamen University, Xiamen, China

[2]School of Public Health, Texas A&M University, College Station, TX, United States

Corresponding Author:
Yan Alicia Hong, PhD
School of Public Health
Texas A&M University
1266 TAMU
College Station, TX, 77840
United States
Phone: 1 9794369343
Fax: 1 9794584267
Email: yhong@sph.tamhsc.edu

Abstract

In October 2015, the "Chinese American Physicians E-Hospital" celebrated its "grand opening" online. All physicians affiliated with this E-Hospital are bilingual Chinese American physicians, who provide services ranging from initial teleconsulting to international transfer and treatment in the United States. Such telemedicine platform for medical tourism not only saves the patients from the hassles of identifying and connecting with an appropriate health service provider but also minimizes the language and cultural barriers. As a growing number of patients from middle- and low-income countries travel to the United States (US) for medical care, we face promising opportunities as well as mounting challenges. The Centers for Disease Control (CDC) in the US has guidance for Americans seeking care overseas, but is not available for international patients seeking care in US. This article opens a dialogue on the challenges associated with flourishing medical tourism and telemedicine, including quality assessment, risk communication, ethical guidelines, and legal concerns.

(J Med Internet Res 2016;18(5):e115) doi:10.2196/jmir.5432

KEYWORDS

E-hospital; medical tourism; telemedicine; ethics

In October 2015, the "Chinese American Physicians E-Hospital" celebrated its "grand opening." Designed specifically for the Chinese population, this E-Hospital launched a website only in Chinese and offers toll free numbers for Chinese patients in both North America and China; it also provides online customer service for Chinese people around the world. All physicians affiliated with this E-Hospital are bilingual Chinese American physicians, who are board certified in the US. These highly qualified and culturally fit physicians provide initial teleconsulting, joint consultation, and ultimately, international transfer. This telemedicine platform for medical tourism not only saves the patients from the hassles of identifying and connecting with an appropriate health services provider but also eliminates the language and cultural barriers.

Medical tourism is defined as traveling to a foreign country to seek medical care. A 2013 online survey from US reported that 27% of patients had engaged in some form of medical tourism. The number of traveling patients and cost for medical tourism has skyrocketed in the recent years, contributing to a flourishing business of $439 billion [1]. Traditionally, medical tourists travel from high-income countries to middle- and low-income countries to seek comparable or identical care at a lower price. Countries such as China, India, Mexico, and Thailand have long been prominent destinations for Western medical tourists [2].

With a growing wealthy class in the middle-income countries and rapid penetration of the Internet, the "one-way flow" of medical tourists has been gradually replaced with "two-way exchange." More and more patients from middle- and low-income countries travel to the high-income countries for better diagnostic capabilities, state-of-the-art medical technologies, and advanced treatment options that may not be available in their home countries. The prominent hospitals like the Mayo clinic, M.D. Anderson Cancer Center, and Cleveland

Clinic are among the most sought-after clinics and hospitals in the US.

A recent report revealed that the number of Chinese patients seeking medical care in US has increased by 400% from 2004 to 2014, spurring a billion dollar business [3]. A new WHO report showed that China's cancer incidence was in rise and accounted for 21.8% of the global total of cancer diagnoses and 26.9% of world's total cancer deaths [4]. Currently, the 5-year cancer survivor rate in China is about 30%, compared to 66% in US. [5]; most clinical trials and new treatments are only available in a selected number of renowned hospitals in US. More than 70% of these Chinese patients seek oncological treatments in US; they typically spend $100,000-$150,000 for their medical trip and pay for their treatment with cash upfront [3]. The aforementioned are the driving forces behind the spurring medical tourism in US [6].

In the wake of high profit in medical tourism, many hospitals are investing to make their services more visible and convenient to their clients from overseas; many hospitals have a department dedicated to provide services to these international patients [3,6]. As the US hospitals and health systems invest in population health initiatives and commit to new payment models that align incentives for keeping patients away from the hospital and clinics, these medical tourists seems to be a logical strategy for keeping hospital beds filled with a new type of fee-for-service customers. E-hospital certainly serves as an efficient platform to attract more medical tourists, and we expect to see more international patients in American hospitals. Accompanying the promising opportunities are the mounting challenges associated with growing number of medical tourists in US.

The first challenge is a lack of reliable assessment of healthcare quality across borders, including a scale of price to outcome. Increasing complexity of healthcare provision and skyrocketing cost of medical care requires more transparency in pricing and standardization of quality assessment. For international patients, a reliable assessment system comparable across borders is especially needed. As Shaw advocated, we need to standardize healthcare standards, and efforts are required in statutory regulation, institutional licensing and accreditation, and increasing transparency of healthcare pricing [7].

Second, there is insufficient communication on the risks associated with medical tourism. The CDC has guidance for Americans seeking care overseas [8], but similar guidance is not available to international patients seeking care in US. Analysis of 91 medical tourism brokers' websites in North America found that most websites failed to report any procedural, postoperative, or legal concerns associated with their services [9]. Analysis of media discourses on medical tourism indicated that risks were less communicated than benefits [10]. Regulations on these medical tourism broker services including E-Hospitals are lagging behind the rapid development of telemedicine and are urgently needed.

Third, there are inadequate policies on ethical concerns related to clinical trials. International patients, traveling thousands of miles to seek treatment in US, are often charged with full emotion and high hope; many are recruited into clinical trials for new but risky treatments [10]. How to address the language and cultural barriers for these patients and how to ensure they understand the potential risks associated with clinical trials have not been addressed in current rules and policies.

Fourth, growing legal concerns are associated with medical tourists. Parallel to the growing number of medical tourists are the more frequent reports of lawsuits associated with medial tourism. Police aids of "maternity hotels" targeting Chinese women giving birth in U.S. and big-dollar lawsuits associated with malpractice in treating international patients have often made headlines. Some of these incidents were attributed to limited communication of risks and survivor rates. Many "gray zones" in medical tourism have not been addressed in our current legislature but need immediate attention.

Rapid growth of medical tourism mirrors accelerative globalization. As Crisp vividly describes "turning the world upside down," ubiquitous Internet access has enabled patients around the world to seek the best care available and facilitates efficient communication of medicine globally [11]. But how to ensure the patient safety, quality of care, ethical issues, and legal concerns remain inadequately addressed, for patients, healthcare providers, and policy makers across borders. Data are the very foundation for the making evidence-based policies and practices; but in our review of literature on medical tourism, we found a dearth of such data. As we enter a new frontier of telemedicine for the old business of medical tourism, we need more research and dialogue on the issues and impact associated with the evolving models of medical consumerism.

Acknowledgments

The open access publishing fees for this article have been sponsored by the Texas A&M University Online Access to Knowledge Fund (OAKFund), supported by the University Libraries and the Office of the Vice President for Research.

Conflicts of Interest

None declared.

References

1. American Medical Tourism Association. Research/Survey/Statistics URL: http://www.medicaltourismassociation.com/en/research-and-surveys.html [accessed 2015-12-10] [WebCite Cache ID 6dgGGyNJ5]

2. Hall CM, editor. Medical Tourism: The Ethics, Regulation, and Marketing of Health Mobility. In: Routledge. New York: Routledge; 2012:1-223.
3. The Wall Street Journal. Desperate Chinese seek medical care abroad URL: http://www.wsj.com/articles/desperate-chinese-seek-medical-care-abroad-1409959941 [accessed 2016-01-08] [WebCite Cache ID 6eN7mOr3a]
4. World Health Organization. 2014. Country Cancer Profile, China URL: http://www.who.int/cancer/country-profiles/chn_en.pdf?ua=1 [accessed 2016-01-08] [WebCite Cache ID 6eN8Db4rT]
5. Zeng H, Zheng R, Guo Y, Zhang S, Zou X, Wang N, et al. Cancer survival in China, 2003-2005: a population-based study. Int J Cancer 2015 Apr 15;136(8):1921-1930 [FREE Full text] [doi: 10.1002/ijc.29227] [Medline: 25242378]
6. Chang J. Chinese spur medical-tourism growth in the US.: August 7, 2015 (China Daily) URL: http://usa.chinadaily.com.cn/epaper/2015-08/07/content_21527300.htm [accessed 2016-01-07] [WebCite Cache ID 6eN7vW8M3]
7. Shaw, CD. How can healthcare standards be standardised? BMJ Quality & Safety 2015 Oct 24;24(10):615-619.
8. Gaines J, Nyugen D. CDC Yellow Book information on Medical Tourism. 2014. Medical Tourism URL: http://wwwnc.cdc.gov/travel/yellowbook/2016/the-pre-travel-consultation/medical-tourism [WebCite Cache ID 6dgGA5Ny3]
9. Jun J, Oh KM. Framing risks and benefits of medical tourism: a content analysis of medical tourism coverage in Korean American community newspapers. J Health Commun 2015;20(6):720-727. [doi: 10.1080/10810730.2015.1018574] [Medline: 25942506]
10. George, Henthorne T. The Incorporation of Telemedicine With Medical Tourism: A Study of Consequences. Journal of Hospitality Marketing & Management 2009 Jun 09;18(5):512-522. [doi: 10.1080/19368620902950097]
11. Crisp N. Turning the World Upside Down: The Search for Global Health in the 21st Century. New York: CRC Press; Jan 15, 2010.

Edited by G Eysenbach; submitted 11.12.15; peer-reviewed by Y Zhang, V Fan; comments to author 31.12.15; revised version received 07.01.16; accepted 21.02.16; published 23.05.16.

Please cite as:
Hong YA
Medical Tourism and Telemedicine: A New Frontier of an Old Business
J Med Internet Res 2016;18(5):e115
URL: http://www.jmir.org/2016/5/e115/
doi:10.2196/jmir.5432
PMID:27215230

Original Paper

Testing the Efficacy of OurSpace, a Brief, Group Dynamics-Based Physical Activity Intervention: A Randomized Controlled Trial

Brandon Irwin[1], PhD; Daniel Kurz[1], MPH; Patrice Chalin[2], PhD; Nicholas Thompson[1], BSc

[1]Digital Physical Activity Laboratory, Department of Kinesiology, Kansas State University, Manhattan, KS, United States

[2]Dependable Software Research Group, Santa Cruz, CA, United States

Corresponding Author:
Brandon Irwin, PhD
Digital Physical Activity Laboratory
Department of Kinesiology
Kansas State University
1A Natatorium
Manhattan, KS, 66506
United States
Phone: 1 7855320660
Fax: 1 5346486
Email: bcirwin@k-state.edu

Abstract

Background: Emerging technologies (ie, mobile phones, Internet) may be effective tools for promoting physical activity (PA). However, few interventions have provided effective means to enhance social support through these platforms. Face-to-face programs that use group dynamics-based principles of behavior change have been shown to be highly effective in enhancing social support through promoting group cohesion and PA, but to date, no studies have examined their effects in Web-based programs.

Objective: The aim was to explore proof of concept and test the efficacy of a brief, online group dynamics-based intervention on PA in a controlled experiment. We expected that the impact of the intervention on PA would be moderated by perceptions of cohesion and the partner's degree of presence in the online media.

Methods: Participants (n=135) were randomized into same-sex dyads and randomly assigned to one of four experimental conditions: standard social support (standard), group dynamics-based–high presence, group dynamics-based–low presence, or individual control. Participants performed two sets of planking exercises (pre-post). Between sets, participants in partnered conditions interacted with a virtual partner using either a standard social support app or a group dynamics-based app (group dynamics-based–low presence and group dynamics-based–high presence), the latter of which they participated in a series of online team-building exercises. Individual participants were given an equivalent rest period between sets. To increase presence during the second set, participants in the group dynamics-based–high presence group saw a live video stream of their partner exercising. Perceptions of cohesion were measured using a modified PA Group Environment Questionnaire. Physical activity was calculated as the time persisted during set 2 after controlling for persistence in set 1.

Results: Perceptions of cohesion were higher in the group dynamics-based–low presence (overall mean 5.81, SD 1.04) condition compared to the standard (overall mean 5.04, SD 0.81) conditions (P=.006), but did not differ between group dynamics-based–low presence and group dynamics-based–high presence (overall mean 5.42, SD 1.07) conditions (P=.25). Physical activity was higher in the high presence condition (mean 64.48, SD 20.19, P=.01) than all other conditions (mean 53.3, SD 17.35).

Conclusions: A brief, online group dynamics-based intervention may be an effective method of improving group cohesion in virtual PA groups. However, it may be insufficient on its own to improve PA.

(J Med Internet Res 2016;18(4):e87) doi:10.2196/jmir.5342

KEYWORDS

physical activity; group dynamics; social media; cohesion; Internet; social support

Introduction

Despite the broad health benefits of physical activity (PA) [1], only approximately 5% of US citizens are actually meeting PA recommendations [2]. The US Centers for Disease Control and Prevention have identified and advocate several evidence-based approaches to promoting PA, including social support-based interventions [3]. Among social support-based interventions are those that involve peer groups [4] and other group-based approaches [5]. Meta-analysis data show that highly effective group interventions are those that include group dynamics-based activities [5]. Group dynamics-based interventions include team-building activities (eg, group goal setting) to facilitate group member interactions with the ultimate goal of enhancing group cohesion [6]. In comparison to delivering interventions to collections of people and individual-based PA programs, group dynamics-based programs have been associated with higher levels of moderate-to-vigorous physical activity (MVPA), program adherence, and levels of social interaction [5]. However, there exist several drawbacks to the group dynamics-based approach to PA promotion—groups are often required to meet in person, which may restrict participation due to geographic location and requires participants to coordinate meeting times amid busy schedules. Additionally, there is a burden placed on staff and practitioners to manage and facilitate the group activities. Thus, strategies that overcome these challenges may help optimize group dynamics-based interventions and free resources to allow for broader reach and effectiveness.

One such strategy might involve the use of the Internet and, in particular, virtual teams [7]. The Internet provides a unique potential for a vast population of people, both sedentary and active, to seek out health information and/or support for behavior change. Additionally, Internet-based tools (eg, social media) can and should facilitate group interactions [8,9], thus decreasing the burden of staff/practitioners. This potential has not gone unnoticed because Internet-based interventions are now being used more often for promoting positive health behaviors, such as PA [10-12]. However, although significant, the overall effects of Internet-delivered interventions focusing on PA promotion have been small [13] and are prone to a variety of drawbacks. For example, participant attrition in Internet-based weight loss programs is typically high (>25%) and those who adhere to the programs often have reduced engagement over time [14]. Given the effectiveness of group dynamics-based programs to impact MVPA and enhance program adherence as well as the ability of the Internet to overcome traditional barriers associated with face-to-face interventions, a sensible strategy for improving Internet-based interventions would be to translate group dynamics-based practices into an Internet-based intervention.

In this study, we developed a brief, online group dynamics-based PA intervention to lead users through a series of virtual team-building activities. The app was designed according to Carron and Spink's [15] team-building model and targeted several key aspects of cohesion development, including group environment, group structure, and group processes. Because many existing Internet-based interventions include peer social support components, we compared the effect of a group dynamics-based intervention with that of a common social support app (ie, a moderated discussion board). We also wished to test for the effect of the intervention on known correlates and mediators of PA, including enjoyment [16,17], perceived exertion [18], and self-determined motivation [19,20].

The primary aim of this study was to test the efficacy of a brief, online group dynamics-based intervention in increasing PA. The secondary aims were to test the moderating effects of group cohesion and the presence of a partner. We created and randomized participants into one of four conditions (individual, standard, group dynamics-based–low presence, group dynamics-based–high presence) to test the following hypotheses:

1. Participants in a brief, group dynamics-based online intervention (group dynamics-based–high presence, group dynamics-based–low presence) will be more physically active than those in a standard social support intervention (standard) after controlling for baseline PA.
2. The impact of the intervention on PA will be moderated by perceptions of group cohesion (group dynamics-based>standard).
3. The impact of the intervention on PA would be moderated by the partners' degree of presence (group dynamics-based–high presence>group dynamics-based–low presence).

Methods

Recruitment

Participants (N=135; 66 males, 69 females; mean age 19.54, SD 1.81 years) were recruited from an introductory level kinesiology course at a large Midwestern university to participate in a single session of a 1-hour "video game" study. Data were collected between March and December 2014. All participants were screened for health risks using the PA Readiness Questionnaire [21] and were awarded course credit for the completion of the study. An alternate assignment for credit was available for the students who did not participate in the research study. Ethical approval for the study was granted by the University's Institutional Review Board (IRB #6318.1).

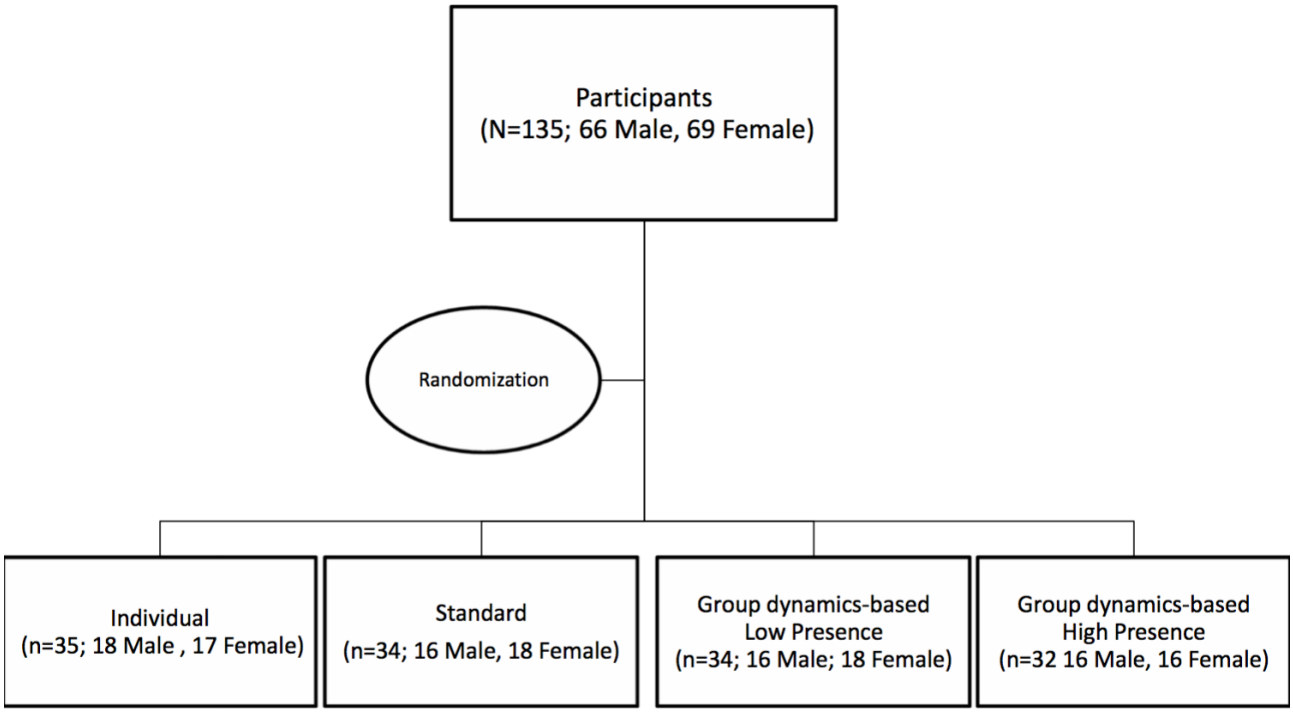
Design

Similar to previous studies testing the impact of group dynamics on PA [22-24], we used a brief intervention in a laboratory setting to test our hypotheses. There are a number of advantages to this approach, including control of extraneous variables and the ability to test the efficacy of the core features of an intervention prototype at minimal financial and human cost. The present study used a randomized 2 (gender) × 4 (condition) × 2 (block) experimental design with repeated measures on the last factor. Each block consisted of an identical series of five planking exercises: front plank, side plank (left), one-leg plank (left), side plank (right), and one-leg plank (right).

Participants were randomly selected from a participant pool and asked to provide times and dates they were available to participate. Participants were then, unknowingly, assigned to same-sex dyads based on their availability. Dyad members were scheduled to participate in the study concurrently. Individual dyad members were sent to separate testing rooms to avoid any

interactions outside of the experiment. Three individuals were unable to be scheduled into a dyad; their results were included in the individual condition. On arrival to the laboratory, dyads were randomly assigned to a condition: individual (n=32, 16 dyads), standard social support (n=34, 17 dyads), group dynamics-based–low presence (n=34, 17 dyads), or group dynamics-based–high presence (n=32, 16 dyads) (Figure 1). In the case that a dyad member failed to appear for their session, the present member was told by the experimenter that "we are experiencing technical difficulties and unable to run the trial today" and was rescheduled for a future time slot (n=17; 14 males, 3 females).

Figure 1. Flowchart of participant distribution.



Procedures

Participants arrived individually, signed an informed consent form, and were instructed to sit in an isolated room—separated from the experimenter and their partner—in front of a computer and began watching a video tutorial that included instructions for their participation during the experiment. Participants were given instructions for proper technique for a series of abdominal planking exercises that they would be performing during the experiment. A computer-generated trainer demonstrated the exercises during both the instructional video and during each block of exercise. Participants were instructed to hold each planking exercise for as long as they could without causing any undue discomfort or pain to themselves.

To minimize the risk of partners becoming aware of the other's proximity, all participants were instructed to wear a pair of noise-canceling headphones for the full duration of the experiment. The headphones doubled as speakers for the computers. Participants were further instructed that if they needed assistance or had questions they should use a chat box provided on their computer, which directly linked them to the experimenter, in lieu of trying to verbally communicate.

Once participants completed the video tutorial, they were instructed to sit on an exercise mat, wait for the virtual trainer to start the exercise, and follow along with that trainer during each exercise (block 1). Once both dyad members were ready to begin, the experimenter initiated the virtual trainer and participants completed the first series of exercises independently and unaware of their partner. All participants performed the planks in the same order with a short (40 sec) rest period between each plank. During each planking exercise, participants were shown a live stream video of themselves exercising, allowing them to check their form against the virtual trainer. This constituted the first block of exercises (block 1).

At the end of block 1, all participants were asked to return to their computer and wait for further instructions. For participants in the individual condition, participants were given a 15-minute rest period in which they were told the average duration they held the planking exercises for, were asked to fill out a brief

self-efficacy survey, and given generic reading material to occupy their time until block 2 began. In any of the partner conditions (standard, group dynamics-based–low presence, or group dynamics-based–high presence), participants were introduced to their partner through the Web app. Participants in a group dynamics-based condition received the full group dynamics-based app, whereas participants in the standard condition received a modified version of the app that removed the majority of team-building activities found in the full version. The standard version of the app was intended to mimic the features found in standard social support apps (eg, a discussion board), where communication is limited to text chat and minimally facilitated (eg, through prompts). In both versions, participants were given the following team task: "The two of you will be performing together as a team. Your team's task is to hold the exercise for as long as possible. Your team's time will be the total number of seconds that your team holds the exercises." Block 2 began following the completion of the group dynamics-based app (when applicable) and a brief rest period.

For block 2, the individual condition followed the same procedures as block 1 whereas the group dynamics-based–low presence and standard conditions followed the same procedures as block 1 except they were now aware they had a partner and were given the aforementioned team task. Participants in the group dynamics-based–high presence condition followed the same procedure as the group dynamics-based–low presence and standard conditions; however, instead of seeing the live video stream of themselves exercising, they were shown the live stream of their partner instead (video streams were blurred to protect participants' confidentiality). This set of planking exercises constituted the second and final block of exercises (block 2).

Following block 2, all participants returned to their computers to complete the final intrinsic motivation survey. Participants in partnered conditions also completed the cohesion questionnaire. Once completed, participants were thanked, debriefed, asked to not discuss the study with their classmates, and dismissed separately to avoid partners meeting each other in person.

Group Dynamics-Based Intervention Description

The group dynamics-based intervention was a theory-based Web app ("OurSpace") informed by Carron and Spink's [15] team-building model. The app included several team-building activities to target different aspects of group dynamics, all of which are evidence-based practices for promoting group cohesion. We developed two versions of the app: a full version that included all group dynamics-based activities and a modified version that included only two activities (intended to mimic traditional social support tools commonly found in Internet-based interventions). In the full version of the app, participants entered their personal information, including first name (although their partner would see a pseudonym), gender, year in school, career goals, and something interesting about themselves, and selected an avatar from a list of generic preset characters. On the following page, each participant was asked to share something they struggled with during the exercises. Partners then exchanged advice on how the other could overcome their struggles (social support). Next, group distinctiveness was established by having the partners vote on and select a team icon and team name. Next, partners worked to solve a simple team-based puzzle together. Completion of the puzzle required partners to cooperatively control an onscreen character using directional arrows. One dyad member was given control of the character up/down movements, whereas the other controlled the right/left movement; coordination and cooperation were required to complete the task. Partners then established a group norm of what they believe the group's expected effort level should be, individually and collectively, agreeing on the expected group effort value using a 1-10 scale. Finally, individual positions within the group were established by telling each dyad member how long they held each exercise and how long their partner held each exercise during block 1. The modified version of the app concluded after the social support page (app descriptions can be found in Table 1). The full version of the app was used by both group dynamics-based conditions, the modified version was used by the standard condition, and individuals did not use the app. Following the completion of the online session, participants reported the task-predicted performance survey and waited to begin the block 2 exercises.

Table 1. Description of the OurSpace app.

Theoretical construct and app feature	Group dynamics-based app	Standard app
Group environment		
Share personal information	Yes	Yes
Team name and icon	Yes	No
Group structure		
Establish group exercise norms	Yes	No
Establish positions within group	Yes	No
Group process		
Team-based puzzle	Yes	No
Social support		
Prompts to provide and receive support	Yes	Yes

Measures/Outcomes

Physical Activity

Physical activity was operationally defined as the amount of time (in seconds) that participants persisted during block 2 after controlling for individual differences in ability and PA during block 1. The summed mean of the time spent performing the five planking exercises constituted a block score. Digital stopwatches were used to measure time spent in each exercise. Time was measured from the moment participants got into position for the exercise until the participant quit the exercise. The rest time between exercises was also calculated. This process was repeated for each plank exercise.

Perceptions of Cohesion

Participant's perception of cohesion was measured using a modified Physical Activity Group Environment Questionnaire (PAGEQ) [25]. Original PAGEQ questions were modified to fit the context of the present study (eg, PAGEQ: "members of our PA group often socialize during exercise time" was modified to "members of our exercise group often socialized during time spent online"). Three items from the original PAGEQ items were omitted from the modified version due to lack of relevance within this study.

The modified PAGEQ measured participants' perceived cohesion based on four dimensions: (1) attraction to group-task (ie, "I like the exercise done in this group"), (2) attraction to group-social (ie, "I enjoyed my social interactions within this online exercise group"), (3) group integration-task (ie, "our group is united in its beliefs about the benefits of the exercises offered in this program"), and (4) group integration-social ("members of our group would likely spend time together after the program ends"). Consistent with the original PAGEQ, each question was answered using a 9-point Likert scale, (ie, 1=very strongly disagree, 5=neither agree nor disagree, and 9=very strongly agree) [18]. Cronbach alpha was used to determine internal consistency reliability; scores for attraction to group-task (Cronbach alpha=.74), attraction to group-social (Cronbach alpha=.85), group integration-task (Cronbach alpha=.80), and group integration-social (Cronbach alpha=.76) were deemed acceptable.

Rating of Perceived Exertion

Ratings of perceived exertion (RPE) were measured using a 10-point RPE scale [26]. Scale measures ranged from 1 meaning "no exertion at all" to 10 meaning "maximal exertion." Participants recorded their own RPEs on a sheet provided to them during the rest period immediately after completing each planking exercise. Scores were calculated as the mean reported RPE for each block.

Task Enjoyment

Task enjoyment was measured using a short 8-item (Cronbach alpha=.74) version of the PA Enjoyment Scale (PAES) [27,28]. Each item was rated on a 7-point bipolar scale beginning with the stem "Please rate how you feel at the moment about the PA you have been doing according to the following scales" (eg, 1=I loved it; 7=I hated it). Previous studies have shown high correlations with the longer 18-item scale (r=.94) [29] and strong reliability (Cronbach alpha=.91) [30].

Motivation

Motivation was measured with the Situational Motivation Scale (SIMS) [31]. The SIMS contained 16 items, which reflected different reasons a participant might be motivated to participate in the exercises. Each item was rated on a 7-point bipolar scale beginning with the stem "Please indicate the answer that best describes the reason why you are currently engaged in the abdominal exercises you are performing. Answer each item according to the following scale" (eg, 1=corresponds not at all; 7=corresponds exactly). There were four subscales (4 items each) based on Self-Determination Theory: amotivation (Cronbach alpha=.94), external regulation (Cronbach alpha=.95), identified regulation (Cronbach alpha=.95), and intrinsic motivation (Cronbach alpha=.93).

Sample Power

An a priori power analysis following F index recommendations indicated that a sample size of n = 32 per condition would be sufficient for detecting a moderate (F=0 . 25) effect with probability >.80. Effect size was determined by a power analysis based on the findings of similar studies [22,23] using G-power software.

Statistical Analyses

Hypothesis Testing

To test the hypotheses that (1) participants using a group dynamics-based app would have higher PA rates than those using the standard app and that (2) PA rates of participants using a group dynamics-based app would be moderated by levels of presence, a 4 (condition) × 2 (gender) ANCOVA was run with block 1 PA, baseline self-efficacy, and measures of intrinsic motivation as the covariates with block 2 PA as the dependent variable. To test the hypothesis that cohesion moderates the impact of PA rates a 3 (condition: all standard, group dynamics-based–low presence, group dynamics-based–high presence) × 2 (gender) MANOVA with each subscale of the modified PAGEQ as dependent variables.

Ancillary Analyses

The RPE (measured at the end of blocks 1 and 2) was analyzed with a 2-way condition × gender ANCOVA with block 1 RPE used as a covariate. Enjoyment (measured after block 2) was analyzed with a 2-way condition × gender ANOVA. Self-efficacy and all motivation subscales (measured at baseline, after block 1, and after block 2) were analyzed with separate 2-way condition × gender repeated measures ANCOVAs with baseline measures used as covariates. All analyses were conducted using SPSS 22 statistical software.

Table 3. Group cohesion by condition.

Cohesion variable and condition vs comparison condition	n	Mean (SD)[a]	*P*
Attraction to group-task			
Standard	34	5.71 (1.37)	
Group dynamics-based–low presence			.08
Group dynamics-based–high presence			.39
Group dynamics-based–low presence	34	6.40 (1.08)	
Standard			.08
Group dynamics-based–high presence			.68
Group dynamics-based–high presence	32	6.13 (1.21)	
Standard			.39
Group dynamics-based–low presence			.68
Attraction to group-social			
Standard	34	4.88 (1.19)	
Group dynamics-based–low presence			.001
Group dynamics-based–high presence			.17
Group dynamics-based–low presence	34	6.06 (1.28)	
Standard			.001
Group dynamics-based–high presence			.19
Group dynamics-based–high presence	32	5.48 (1.33)	
Standard			.17
Group dynamics-based–low presence			.19
Group integration-task			
Standard	34	4.12 (1.18)	
Group dynamics-based–low presence			<.001
Group dynamics-based–high presence			.02
Group dynamics-based–low presence	34	5.20 (1.04)	
Standard			<.001
Group dynamics-based–high presence			.43
Group dynamics-based–high presence	32	4.86 (0.93)	
Standard			.02
Group dynamics-based–low presence			.43
Group integration-social			
Standard	34	4.61 (1.64)	
Group dynamics-based–low presence			.04
Group dynamics-based–high presence			.48
Group dynamics-based–low presence	34	5.46 (1.22)	
Standard			.04
Group dynamics-based–high presence			.44
Group dynamics-based–high presence	32	5.02 (1.17)	
Standard			.48
Group dynamics-based–low presence			.44

[a]Perceived cohesion scales ranged from 1 to 9.

Results

Sample Population

The total sample consisted of 135 college-aged participants (66 males, 69 females; age mean 19.54, SD 1.81). No participants dropped out of the study before completing their sessions.

Preliminary Analyses

An intraclass correlation analysis was run to detect potential agreement or "clustering" of PA and cohesion scores within dyads. Results for perception of cohesion were analyzed according Carron and colleagues' [32] recommendations on determining the degree to which perceptions were shared within groups (or "groupness"). Criteria for detecting a small groupness effect was set at an intraclass correlation coefficient (ICC) of greater than or equal to .40 for attraction to group-social and attraction to group-task and an ICC of greater than or equal to .60 for group integration-social and group integration-task. Under these criteria, there was no evidence of a group clustering for perception of cohesion scores (attraction to group-task: ICC=.258, *P*=.15; attraction to group-social: ICC=.088, *P*=.37; group integration-social: ICC=.505, *P*=.008; group integration-task: ICC=.253, *P*=.16). Results of PA indicated that scores were not clustered in dyads (ICC=.173, *P*=.08) for any conditions (individual: ICC=.268 *P*=.15; standard: ICC=.067, *P*=.40; group dynamics-based–low presence: ICC=–.088, *P*=.64; individual: ICC=.341, *P*=.09). Thus, all following analyses of PA and perceived cohesion were conducted at the individual level.

Physical Activity

Physical activity was significantly greater in the group dynamics-based–high presence condition than all other conditions, ($F_{3,121}$= 3.75, *P*=.01). There were no other significant differences between conditions. There was also a gender main effect; males were more physically active than females ($F_{1,121}$=7.78, *P*=.006). Means of the analysis can be found in Table 2.

Table 2. Physical activity measured as mean persistence by condition and gender.

Condition	Overall			Male[a]		Female[a]	
	n	Mean (SD)	Range	n	Mean (SD)	n	Mean (SD)
Individual	35	50.56 (18.43)	15.6-87.2	18	55.74 (20.32)	17	45.07 (14.85)
Standard	34	54.98 (17.12)	27.4-86.0	16	57.18 (15.32)	18	53.03 (18.80)
Group dynamics-based–low presence	34	54.36 (16.50)	22.0-110.8	16	60.48 (16.98)	18	48.93 (14.41)
Group dynamics-based–high presence	32	64.48 (20.19)[a]	32.8-102.6	16	73.86 (17.07)	16	55.09 (19.05)

[a]Significant at the *P*<.05 level.

Group Cohesion

There was significant main effect for condition ($F_{8,184}$=2.77, *P*=.01). There were significant differences between conditions in all dimensions of cohesion (*P*<.001 to *P*=.04), except for attraction to group-task, (*P*=.06). Overall, where differences existed, mean scores were higher in the group dynamics-based conditions than the standard conditions. Specific differences were identified using a Scheffe post hoc analysis and are identified along with the means and standard deviations in Table 3.

Ancillary Analyses

There were no significant differences in RPE between conditions ($F_{3,127}$=1.83, *P*=.14) or gender ($F_{1,127}$=0.30, *P*=.58), on enjoyment between conditions ($F_{3,127}$=1.60, *P*=.19) or gender ($F_{1,127}$=0.22, *P*=.64), or in motivation between conditions (intrinsic motivation: $F_{3,123}$=0.50, *P*=.01; identified regulation: $F_{3,126}$=0.41, *P*=.74; external regulation: $F_{3,126}$=0.49, *P*=.68; amotivation: $F_{3,126}$= 2.56, *P*=.06) or gender (intrinsic motivation: $F_{1,123}$=0.03, *P*<.001; identified regulation: $F_{1,126}$=0.68, *P*=.41; external regulation: $F_{1,126}$=0.50, *P*=.48; amotivation: $F_{1,126}$=0.13, *P*=.72) over time.

Discussion

The primary aims of this study were to test the efficacy of a brief, online group dynamics-based intervention to increase PA. We also examined the ability of the intervention to impact group cohesion and tested the moderating effect of presence. We hypothesized that the group dynamics-based intervention would impact group cohesion and that higher perceptions of cohesion would be related to higher levels of PA. We also hypothesized that participants who were more visually present to their partner would be more physically active. Our hypotheses were partially supported. Although groups were more physically active than individuals; overall, the only intervention feature that impacted PA was the degree of the partner's presence. Individuals were more active when their partner was virtually present during exercise. Although the intervention impacted the group cohesion mediating process, group cohesion was unrelated to PA. Potential explanations and implications for research and practice are discussed.

Principal Results

Physical Activity

The hypothesis that the brief, online group dynamics-based intervention would impact PA was partially supported. Participants receiving the group dynamics-based intervention

were more physically active than those exercising alone and those receiving a minimal social support intervention, but only when one's partner was highly present. Further, although participants were more active in the high presence condition, our RPE data show that they did not perceive themselves to be working any harder than those in the less physically active conditions, suggesting that exercising with a highly present virtual partner may help overcome barriers related to exercise intensity [33]. Further, all measures of motivation were equal across conditions, suggesting that exercising with a partner under these conditions poses little risk for undermining one's self-determined motivation for PA. This is an encouraging result because self-determined motivation is correlated with long-term maintenance of PA behavior change [20].

Our finding that the intervention did not impact PA even though we impacted the group cohesion mediator contradicts a large body of evidence linking group dynamics-based interventions, group cohesion, and PA [5,34]. However, there are a number of notable differences between the conditions within our brief, online group dynamics-based intervention for PA and previous group dynamics-based interventions for PA, including mode of delivery, dose of intervention, and the type of behavior targeted. First, our intervention was delivered entirely online, whereas previous group dynamics-based interventions have been delivered either face-to-face or partially online [35]. A completely online intervention has the obvious drawback of limiting the amount of communication, presence, and identifiability between participants, all of which are key factors impacting group dynamics and subsequent performance and PA. Despite our efforts in this intervention to enhance such group processes through the Web-based app, they may have failed to meet a minimum threshold for behavior change and may rely on face-to-face interactions to impact PA.

Second, the current intervention targeted PA intensity and duration, whereas many other group dynamics-based interventions target frequency of PA and program adherence. Participants in previous face-to-face group dynamics-based interventions for PA may be more physically active as a result of the intervention because of the intervention impact on frequency of PA bouts and/or attendance at program-related activities (ie, adherence) [5], not because they are more likely to exercise longer or harder within bouts than nonadherers. To date, research has not examined this claim. Third, the importance of the goal given to participants may have undermined their motivation to be physically active. Self-selected goals have been shown to lead to higher goal attainment and effort than non-self-selected goals [33]. In our intervention, groups were assigned the goal of holding the plank exercises for as long as they could. In previous group dynamics-based interventions, groups typically have the ability to select their own PA goals [35].

Group Cohesion

The second hypothesis that groups using the group dynamics-based app would report higher perceptions of cohesion compared to the standard social support app was supported. Results showed that, with the exception of attraction to group-task, perceptions of cohesion were higher for those who participated in the enhanced group dynamics-based intervention compared to those in the comparison condition who only received a minimal dose of social support. This finding is consistent with past face-to-face group dynamics-based studies in which the use of group dynamics-based principles has been shown to improve perceptions of cohesion among groups [6]. This finding is encouraging for online interventions with social support components, suggesting that such components might be fruitfully designed to impact group processes known to mediate the impact of interventions on PA. Further, our data suggest that impacting this mediator can be done rather swiftly, considering the study consisted of a single 1-hour visit during which only 7 to 9 minutes were spent using the group dynamics-based app.

Partner Presence

The third hypothesis, that increased group member presence would moderate PA, was supported. In fact, presence was the only factor to impact PA, in addition to the provision of performance feedback. This corroborates a large body of research showing the impact of virtual presence on group performance, including PA [36,37], but is the first to demonstrate these effects within the context of an online group dynamics-based intervention for PA.

Implications

Our data suggest that impacting cohesion in online groups is feasible using a group dynamics-based app. More generally, this study suggests that online peer-to-peer interactions can be facilitated to impact group-level mediators of PA through purposeful, theory-based design. Given the obstacles and resources needed to implement and participate in face-to-face group-based interventions, further research and optimization of online group dynamics-based interventions such as this is warranted. As mentioned, future research can and should test the impact of an online group dynamics-based intervention on PA over longer durations and bouts of exercise (ie, adherence) and vary and test the dose of online versus face-to-face interaction on PA outcomes.

Limitations

This study has several limitations. First, without a standard app high-presence condition we cannot conclude with certainty whether the higher levels of PA are attributable to the group dynamics-based components, the heightened presence of one's partner, or a synergistic effect of both. Indeed, higher levels of online presence are associated with a variety of positive outcomes (eg, increased productivity, fewer antisocial behaviors, increased PA) and may be operating independently of group cohesion. This is likely the case because presence was manipulated only after participants had completed the team-building activities, not during them, and that all dimensions of cohesion were slightly higher in the group dynamics-based low-presence condition than in the high-presence condition, although these differences were nonsignificant. Future research would do well to untangle the impact of presence in online PA interventions and whether group dynamics-based components add value to such interventions. Second, we used a convenience sample of college students recruited from an introductory level

kinesiology course, who may be more used to and respond differently to digital technology than other populations (eg, adults, elderly). Further, this was a brief intervention conducted in a laboratory with only one bout of exercise. Whether the impact of this intervention on cohesion and PA can be sustained over time and across different real-world settings warrants further study.

Conclusions

This study tested an online group dynamics-based intervention to increase PA. Our brief intervention successfully impacted the mediating process, group cohesion, but did not impact PA. Additionally, virtual presence of group members had no effect on either perceptions of cohesion or PA.

In summary, online group dynamics-based apps may be a practical resource that can be used to overcome traditional barriers to utilizing group dynamics, such as the geographic distance between partners and the burden of staff/practitioners having to facilitate team-building exercises. In addition, this study found that a group dynamics-based tool in an online intervention may provide a more engaging social environment for participants to interact in than those of a standard social support app. Further study is needed to determine whether incorporating group dynamics-based principles into PA interventions of longer duration can impact cohesion and PA in real-world settings.

Acknowledgments

The author would like to acknowledge Chase Therkelsen, Brendan McCluskey, and Morgan Shoemaker for their instrumental role in the acquisition of data, and David Dzewaltoski for his comments on early drafts of the manuscript.

Authors' Contributions

BCI conceived the study, led the study design, and development of the OurSpace app. DLK made intellectual contributions to the study design and OurSpace app, drafted sections of the manuscript, collected a significant portion of the data, and assisted in the statistical analysis. PC made intellectual contributions to the study design, programmed the group dynamics-based app, and made minor edits to the manuscript. NST made intellectual contributions to the study design and made minor edits to the manuscript. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

References

1. US Department of Health and Human Services. Physical Activity and Health: A Report of The Surgeon General. Atlanta, GA: Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion; 1996.
2. Troiano RP, Berrigan D, Dodd KW, Mâsse LC, Tilert T, McDowell M. Physical activity in the United States measured by accelerometer. Med Sci Sports Exerc 2008 Jan;40(1):181-188. [doi: 10.1249/mss.0b013e31815a51b3] [Medline: 18091006]
3. Kahn EB, Ramsey LT, Brownson RC, Heath GW, Howze EH, Powell KE, et al. The effectiveness of interventions to increase physical activity. A systematic review. Am J Prev Med 2002 May;22(4 Suppl):73-107. [Medline: 11985936]
4. Dishman RK, Buckworth J. Increasing physical activity: a quantitative synthesis. Med Sci Sports Exerc 1996 Jun;28(6):706-719. [Medline: 8784759]
5. Burke S, Carron A, Eys M. Group versus individual approach? A meta-analysis of the effectiveness of interventions to promote physical activity. Sport Ex Psy Rev 2006;2(1):19-35.
6. Estabrooks PA, Harden SM, Burke SM. Group dynamics in physical activity promotion: what works? Soc Pers Psychol Comp 2012;6(1):18-40. [doi: 10.1111/j.1751-9004.2011.00409.x]
7. Hertel G, Geister S, Konradt U. Managing virtual teams: a review of current empirical research. Hum Resour Manage Rev 2005;15(1):69-95. [doi: 10.1016/j.hrmr.2005.01.002]
8. Lefebvre R, Bornkessel A. Digital social networks and health. Circulation 2013;127:1829-1836. [doi: 10.1161/CIRCULATIONAHA.112.000897]
9. Li J, Barnett T, Goodman E, Wasserman R, Kemper A, American Heart Association Atherosclerosis, Hypertension and Obesity in the Young Committee of the Council on Cardiovascular Disease in the Young, Council on EpidemiologyPrevention,Council on Nutrition, Physical ActivityMetabolism. Approaches to the prevention and management of childhood obesity: the role of social networks and the use of social media and related electronic technologies: a scientific statement from the American Heart Association. Circulation 2013 Jan 15;127(2):260-267 [FREE Full text] [doi: 10.1161/CIR.0b013e3182756d8e] [Medline: 23212719]
10. Wantland DJ, Portillo CJ, Holzemer WL, Slaughter R, McGhee EM. The effectiveness of Web-based vs. non-Web-based interventions: a meta-analysis of behavioral change outcomes. J Med Internet Res 2004 Nov 10;6(4):e40 [FREE Full text] [doi: 10.2196/jmir.6.4.e40] [Medline: 15631964]

11. van den Berg MH, Schoones JW, Vliet Vlieland TP. Internet-based physical activity interventions: a systematic review of the literature. J Med Internet Res 2007;9(3):e26 [FREE Full text] [doi: 10.2196/jmir.9.3.e26] [Medline: 17942388]
12. Norman GJ, Zabinski MF, Adams MA, Rosenberg DE, Yaroch AL, Atienza AA. A review of eHealth interventions for physical activity and dietary behavior change. Am J Prev Med 2007 Oct;33(4):336-345 [FREE Full text] [doi: 10.1016/j.amepre.2007.05.007] [Medline: 17888860]
13. Davies CA, Spence JC, Vandelanotte C, Caperchione CM, Mummery WK. Meta-analysis of internet-delivered interventions to increase physical activity levels. Int J Behav Nutr Phys Act 2012;9:52 [FREE Full text] [doi: 10.1186/1479-5868-9-52] [Medline: 22546283]
14. Bennett GG, Glasgow RE. The delivery of public health interventions via the Internet: actualizing their potential. Annu Rev Public Health 2009;30:273-392. [doi: 10.1146/annurev.publhealth.031308.100235] [Medline: 19296777]
15. Carron A, Spink K. Team building in an exercise setting. Sport Psychol. 1993. (1) p. 8-18 URL: http://journals.humankinetics.com/tsp-back-issues/tspvolume7issue1march/teambuildinginanexercisesetting [WebCite Cache ID 6hIrCBpqz]
16. Murcia JA, Román ML, Galindo CM, Alonso N, González-Cutre D. Peers' influence on exercise enjoyment: a self-determination theory approach. J Sports Sci Med 2008;7(1):23-31 [FREE Full text] [Medline: 24150130]
17. Fox LD, Rejeski WJ, Gauvin L. Effects of leadership style and group dynamics on enjoyment of physical activity. Am J Health Promot 2000;14(5):277-283. [Medline: 11009853]
18. Oka RK, Gortner SR, Stotts NA, Haskell WL. Predictors of physical activity in patients with chronic heart failure secondary to either ischemic or idiopathic dilated cardiomyopathy. Am J Cardiol 1996 Jan 15;77(2):159-163. [Medline: 8546084]
19. Ryan RM, Patrick H. Self-determination theory and physical activity: the dynamics of motivation in development and wellness. Hell J Psych. 2009. p. 107-124 URL: http://selfdeterminationtheory.org/SDT/documents/2009_RyanWilliamsPatrickDeci_HJOP.pdf [WebCite Cache ID 6hIqLVSIa]
20. Teixeira PJ, Carraça EV, Markland D, Silva MN, Ryan RM. Exercise, physical activity, and self-determination theory: a systematic review. Int J Behav Nutr Phys Act 2012;9:78 [FREE Full text] [doi: 10.1186/1479-5868-9-78] [Medline: 22726453]
21. Cardinal BJ, Esters J, Cardinal MK. Evaluation of the revised physical activity readiness questionnaire in older adults. Med Sci Sports Exerc 1996 Apr;28(4):468-472. [Medline: 8778552]
22. Feltz DL, Kerr NL, Irwin BC. Buddy up: the Köhler effect applied to health games. J Sport Exerc Psychol 2011 Aug;33(4):506-526. [Medline: 21808077]
23. Irwin BC, Scorniaenchi J, Kerr NL, Eisenmann JC, Feltz DL. Aerobic exercise is promoted when individual performance affects the group: a test of the Kohler motivation gain effect. Ann Behav Med 2012 Oct;44(2):151-159. [doi: 10.1007/s12160-012-9367-4] [Medline: 22576339]
24. Weber B, Hertel G. Motivation gains of inferior group members: a meta-analytical review. J Pers Soc Psychol 2007 Dec;93(6):973-993. [doi: 10.1037/0022-3514.93.6.973] [Medline: 18072849]
25. Estabrooks P, Carron A. The Physical Activity Group Environment Questionnaire: an instrument for the assessment of cohesion in exercise classes. Group Dyn-Theor Res. 2000. (3) p. 230-243 URL: http://psycnet.apa.org/psycinfo/2000-12222-003 [WebCite Cache ID 6hIqRV143]
26. Borg GA. Psychophysical bases of perceived exertion. Med Sci Sports Exerc 1982;14(5):377-381. [Medline: 7154893]
27. Paxton RJ, Nigg C, Motl RW, Yamashita M, Chung R, Battista J, et al. Physical Activity Enjoyment Scale short form--does it fit for children? Res Q Exerc Sport 2008 Sep;79(3):423-427. [doi: 10.1080/02701367.2008.10599508] [Medline: 18816956]
28. Kendzierski D, DeCarlo K. Physical Activity Enjoyment Scale: two validation studies. J Sport Exerc Psychol 1991 Mar;13:50-64.
29. Raedeke T. The relationship between enjoyment and affective responses to exercise. J Appl Sport Psychol 2007 Feb 14;19(1):105-115. [doi: 10.1080/10413200601113638]
30. McArthur LH, Raedeke TD. Race and sex differences in college student physical activity correlates. Am J Health Behav 2009;33(1):80-90. [Medline: 18844523]
31. Guay F, Vallerand R, Blanchard C. On the assessment of situational intrinsic and extrinsic motivation: The Situational Motivation Scale (SIMS). Motiv Emotion 2000;24(3):175-213 [FREE Full text] [doi: 10.1023/A:1005614228250]
32. Carron A, Brawley L, Eys M, Bray S, Dorsch K, Estabrooks P. Do individual perceptions of group cohesion reflect shared beliefs? An empirical analysis. Small Grp Rsch 2003 Aug;34(4):468-496. [doi: 10.1177/1046496403254274]
33. Williams DM. Exercise, affect, and adherence: an integrated model and a case for self-paced exercise. J Sport Exerc Psychol 2008 Oct;30(5):471-496 [FREE Full text] [Medline: 18971508]
34. Evans C, Dion K. Group cohesion and performance a meta-analysis. Small Grp Rsch. 2012 Dec. (6) p. 690-701 URL: http://sgr.sagepub.com/content/43/6/690.short [WebCite Cache ID 6hIr4sMsN]
35. Estabrooks PA, Bradshaw M, Dzewaltowski DA, Smith-Ray RL. Determining the impact of Walk Kansas: applying a team-building approach to community physical activity promotion. Ann Behav Med 2008 Aug;36(1):1-12. [doi: 10.1007/s12160-008-9040-0] [Medline: 18607666]
36. Martins LL, Gilson LL, Maynard MT. Virtual teams: what do we know and where do we go from here? Journal of Management. J Manage. 2004 Dec. (6) p. 805-835 URL: http://jom.sagepub.com/content/30/6/805.abstract [WebCite Cache ID 6hIrJ5NYo]

37. Lount R, Park E, Kerr N, Messe L, Seok D. Evaluation concerns and the Köhler effect the impact of physical presence on motivation gains. Small Grp Rsch. 2008 Dec. (6) p. 795-812 URL: http://sgr.sagepub.com/content/39/6/795.short [WebCite Cache ID 6hIrN0huZ]

Abbreviations

ICC: intraclass correlation coefficient
MVPA: moderate-to-vigorous physical activity
PA: physical activity
PAGEQ: Physical Activity Group Environment Questionnaire
RPE: ratings of perceived exertion

Edited by G Eysenbach; submitted 18.11.15; peer-reviewed by D Feltz, R Shaw, S Blair; comments to author 10.12.15; revised version received 04.02.16; accepted 09.02.16; published 06.05.16.

Please cite as:
Irwin B, Kurz D, Chalin P, Thompson N
Testing the Efficacy of OurSpace, a Brief, Group Dynamics-Based Physical Activity Intervention: A Randomized Controlled Trial
J Med Internet Res 2016;18(4):e87
URL: http://www.jmir.org/2016/4/e87/
doi:10.2196/jmir.5342
PMID:27154301

Original Paper

Effectiveness of a Web-Based Screening and Fully Automated Brief Motivational Intervention for Adolescent Substance Use: A Randomized Controlled Trial

Nicolas Arnaud[1*], PhD; Christiane Baldus[1*], PhD; Tobias H Elgán[2], PhD; Nina De Paepe[3], M Sc; Hanne Tønnesen[4,5], MD, PhD; Ladislav Csémy[6], PhD; Rainer Thomasius[1], MD

[1]German Centre for Addiction Research in Childhood and Adolescence (DZSKJ), Centre for Psychosocial Medicine, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

[2]STAD, Centre for Psychiatry Research, Department of Clinical Neuroscience, Karolinska Institutet & Stockholm Health Care Services, Stockholm County Council, Stockholm, Sweden

[3]VAD, Vereniging voor Alcohol- en andere Drugproblemen, Brussels, Belgium

[4]Clinical Health Promotion Centre, Faculty of Medicine, Lund University & Skåne University Hospital, Malmö, Sweden

[5]WHO Collaborating Centre, Bispebjerg & Frederiksberg Hospital Faculty of Medicine, University of Southern Denmark, Copenhagen, Denmark

[6]Centre for Epidemiological and Clinical Research on Addictions, National Institute of Mental Health (NIMH), Klecany, Czech Republic

[*]these authors contributed equally

Corresponding Author:
Nicolas Arnaud, PhD
German Centre for Addiction Research in Childhood and Adolescence (DZSKJ)
Centre for Psychosocial Medicine
University Medical Center Hamburg-Eppendorf
Martinistraße 52
Hamburg, 20246
Germany
Phone: 49 040 7410 59861
Fax: 49 040 7410 56571
Email: n.arnaud@uke.de

Abstract

Background: Mid-to-late adolescence is a critical period for initiation of alcohol and drug problems, which can be reduced by targeted brief motivational interventions. Web-based brief interventions have advantages in terms of acceptability and accessibility and have shown significant reductions of substance use among college students. However, the evidence is sparse among adolescents with at-risk use of alcohol and other drugs.

Objective: This study evaluated the effectiveness of a targeted and fully automated Web-based brief motivational intervention with no face-to-face components on substance use among adolescents screened for at-risk substance use in four European countries.

Methods: In an open-access, purely Web-based randomized controlled trial, a convenience sample of adolescents aged 16-18 years from Sweden, Germany, Belgium, and the Czech Republic was recruited using online and offline methods and screened online for at-risk substance use using the CRAFFT (Car, Relax, Alone, Forget, Friends, Trouble) screening instrument. Participants were randomized to a single session brief motivational intervention group or an assessment-only control group but not blinded. Primary outcome was differences in past month drinking measured by a self-reported AUDIT-C-based index score for drinking frequency, quantity, and frequency of binge drinking with measures collected online at baseline and after 3 months. Secondary outcomes were the AUDIT-C-based separate drinking indicators, illegal drug use, and polydrug use. All outcome analyses were conducted with and without Expectation Maximization (EM) imputation of missing follow-up data.

Results: In total, 2673 adolescents were screened and 1449 (54.2%) participants were randomized to the intervention or control group. After 3 months, 211 adolescents (14.5%) provided follow-up data. Compared to the control group, results from linear mixed models revealed significant reductions in self-reported past-month drinking in favor of the intervention group in both the non-imputed (P=.010) and the EM-imputed sample (P=.022). Secondary analyses revealed a significant effect on drinking frequency (P=.037) and frequency of binge drinking (P=.044) in the non-imputation-based analyses and drinking quantity (P=.021)

when missing data were imputed. Analyses for illegal drug use and polydrug use revealed no significant differences between the study groups (*P*s>.05).

Conclusions: Although the study is limited by a large drop-out, significant between-group effects for alcohol use indicate that targeted brief motivational intervention in a fully automated Web-based format can be effective to reduce drinking and lessen existing substance use service barriers for at-risk drinking European adolescents.

Trial Registration: International Standard Randomized Controlled Trial Registry: ISRCTN95538913; http://www.isrctn.com/ISRCTN95538913 (Archived by WebCite at http://www.webcitation.org/6XkuUEwBx)

(J Med Internet Res 2016;18(5):e103) doi:10.2196/jmir.4643

KEYWORDS

substance use; adolescents; brief intervention; web-based intervention; motivational interviewing; randomized controlled trial

Introduction

Early misuse of alcohol and other drugs is widespread in Europe with higher prevalence compared to other regions in the world such as the United States [1, 2]. Although temporary substance misuse is a common and partly normative phenomenon in youth development [3,4], adolescence is a critical period for the development of addiction problems. This period is typical for initiation, and rapid escalation of individual problematic substance use patterns into clinically significant problems can be observed among a substantial proportion of youth in Europe [5, 6]. Early excessive drinking and combined use of alcohol with other psychoactive substances (ie, polydrug use) are of particular relevance [7-11] due to the associated adverse effects on physical, psychological, and social functioning that put youth at a heightened risk for long-lasting disadvantages [12-14].

The widespread use of alcohol and other drugs suggests that current capacities to prevent youth from initiating alcohol and other drug use are limited [15, 16]. Prevention efforts should therefore target at-risk youth with indicated preventive interventions [17, 18]. Effective methods to prevent risky substance use and addiction problems are in principle available, but existing health service provision is limited in accessibility and acceptability [19, 20] with the result that interventions are often provided too late and do not reach the majority of at-risk subjects [21, 22].

Web-based intervention programs have been increasingly acknowledged in their capacity to lessen existing service barriers particularly for at-risk populations [23-25]. Moreover, fully automatic delivery (ie, stand-alone or self-guided with no clinician involvement) allow for standardized delivery and can be disseminated cost-effectively at a large scale [24]. Due to the high Internet access rates in contemporary societies and the fact that youth typically use the Internet when searching for information about alcohol and drugs and also are reluctant to disclose alcohol- and drug-related behavior in face-to-face contacts, Web-based interventions hold promise for younger populations [26, 27]. Evidence indicates that fully automated brief motivational interventions can reduce drinking and related harms for emerging adult at-risk drinkers up to 12 months after the intervention [28-31].

The literature on Web-based interventions for illegal drug use is not as developed as it is for alcohol, but a recent meta-analysis (including 10 studies) suggests that overall the effects are somewhat smaller (g=0.16) compared to drinking (g=0.20-0.39) but significant and (as for alcohol interventions) independent of intervention venue (home vs research setting) and level of guidance through the intervention [32-34].

Although previous studies that have proven the usefulness of Web-based motivational interventions to address substance use and related problems mainly targeted emerging adults [35], the motivational methods that have been studied and found effective are relevant for many risk factors in adolescence, such as their susceptibility to peer influences [36-39]. Motivational interventions are based on the therapeutic style and techniques put forward by Motivational Interviewing (MI) [40], which makes a strong case for conceptual compatibility with adolescent-specific needs for autonomy, subjective perceptions of invulnerability to apparent health risks, and appraisal for short-term benefits at the cost of possible long-term adverse effects [17, 37, 38]. Although effects of MI interventions for alcohol misuse and applied techniques in brief interventions such as normative feedback have recently been summarized as rather small [41, 42], relevant previous studies have shown favorable results from face-to-face [43, 44] as well as Web-based MI-interventions among young adults [29-31]. However, studies targeting mid-to-late adolescents (aged 16-18 years) are lacking [45] despite the fact that this is a critical period for establishing problematic alcohol and other drug use [6, 46] and rapid acceleration for first use of illegal drugs [47, 48].

The purpose of this study was therefore to test the effectiveness of a fully automated Web-based brief MI in a sample of at-risk substance-using adolescents in four European countries. Our primary hypothesis was that participants in the intervention group would report significantly lower levels of past-month drinking (frequency, quantity, and frequency of binge drinking) at 3-month follow-up relative to baseline when compared to an assessment-only control group. Additional hypotheses concerned differences in past-month illegal drug use and combined use of alcohol and illegal drugs.

Methods

A two-armed multisite randomized controlled trial (RCT) design was applied in Sweden, Belgium, the Czech Republic, and Germany. Inclusion criteria were being 16-18 years old, online access, informed consent, and a positive screening for at-risk substance use. Baseline assessment was collected at study entry and a follow-up assessment was collected 3 months after

baseline assessment. Figure 1 displays the trial design (see Multimedia Appendix 1[49] for the CONSORT EHEALTH checklist).

Five university research centers in Europe developed the purely Web-based content of the WISEteens portal between June 2011 and March 2012. The IT platform was established together with GAIA AG, Hamburg. The landing page (see Figure 2) was designed to create an appealing first impression using visual material (eg, pictures, video). It described the main features of the study by highlighting confidentiality, content and source credibility, and provided a brief guided enrollment procedure [50]. Key to developing the content was the integration of MI principles and techniques in a single session together with an open-access delivery format applying a design to match end-user characteristics and preferences. Furthermore, the content should be acceptable, easy to use, and perceived as relevant by the target group [50-53].

All material was developed in a multidisciplinary team including experts in clinical health promotion, developmental and clinical psychology, and certified behavioral therapists. It was first developed in English and then translated by professional offices into the respective countries' languages. The Web portal was simultaneously launched in all four countries in June 2012 with recruitment until March 2013. Ethical approval was granted by the responsible Ethics Committees in all participating countries: Chamber of Physicians Hamburg (Germany), Prague Psychiatric Centre (Czech Republic), University Hospital of Antwerp and the University of Antwerp (Belgium), and the Regional Ethics Board in Stockholm (Sweden). The trial design was published [54], and the study was registered in a public database. No content or methodological modifications were made after trial commencement.

Figure 1. Participant flow.

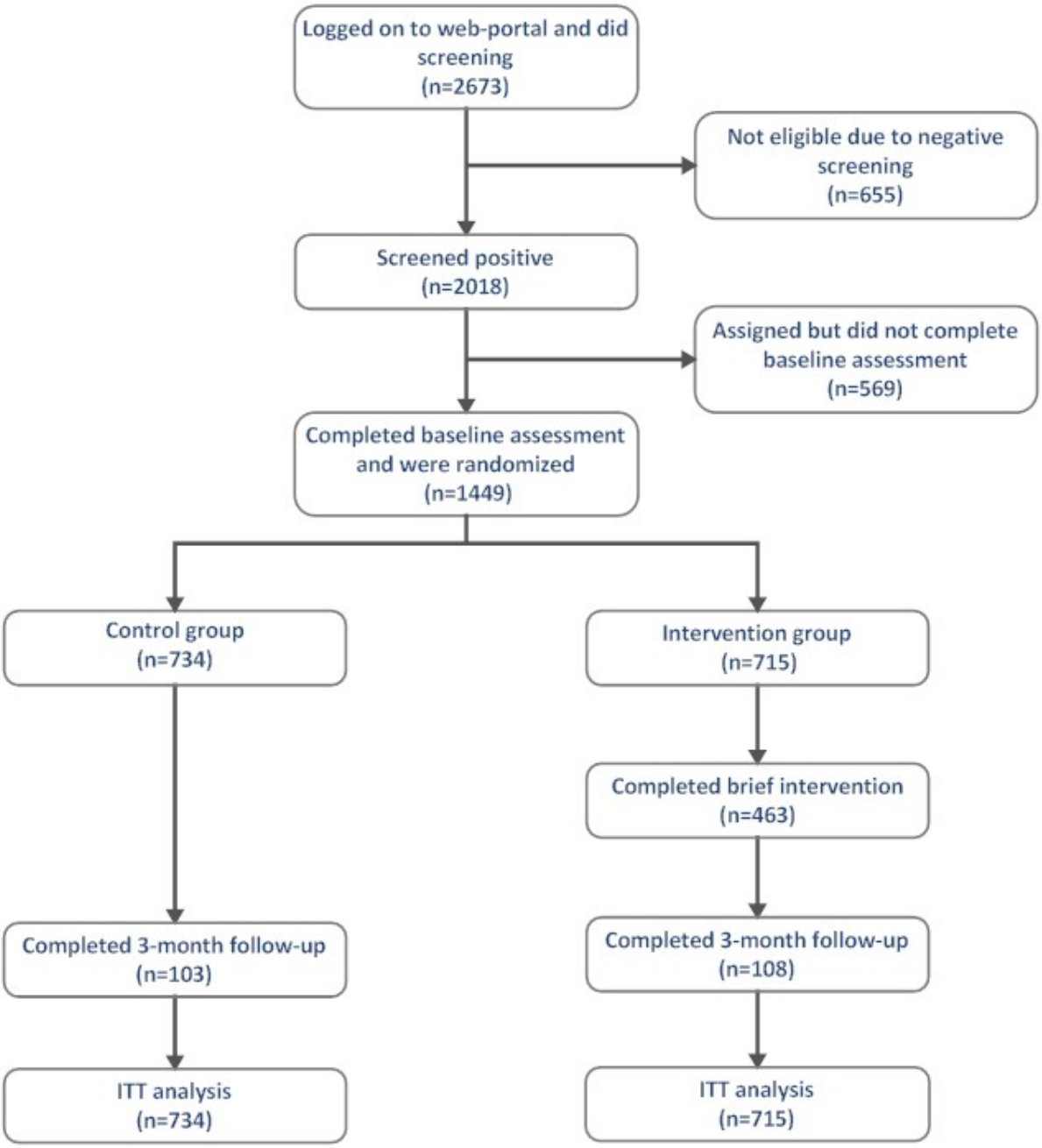
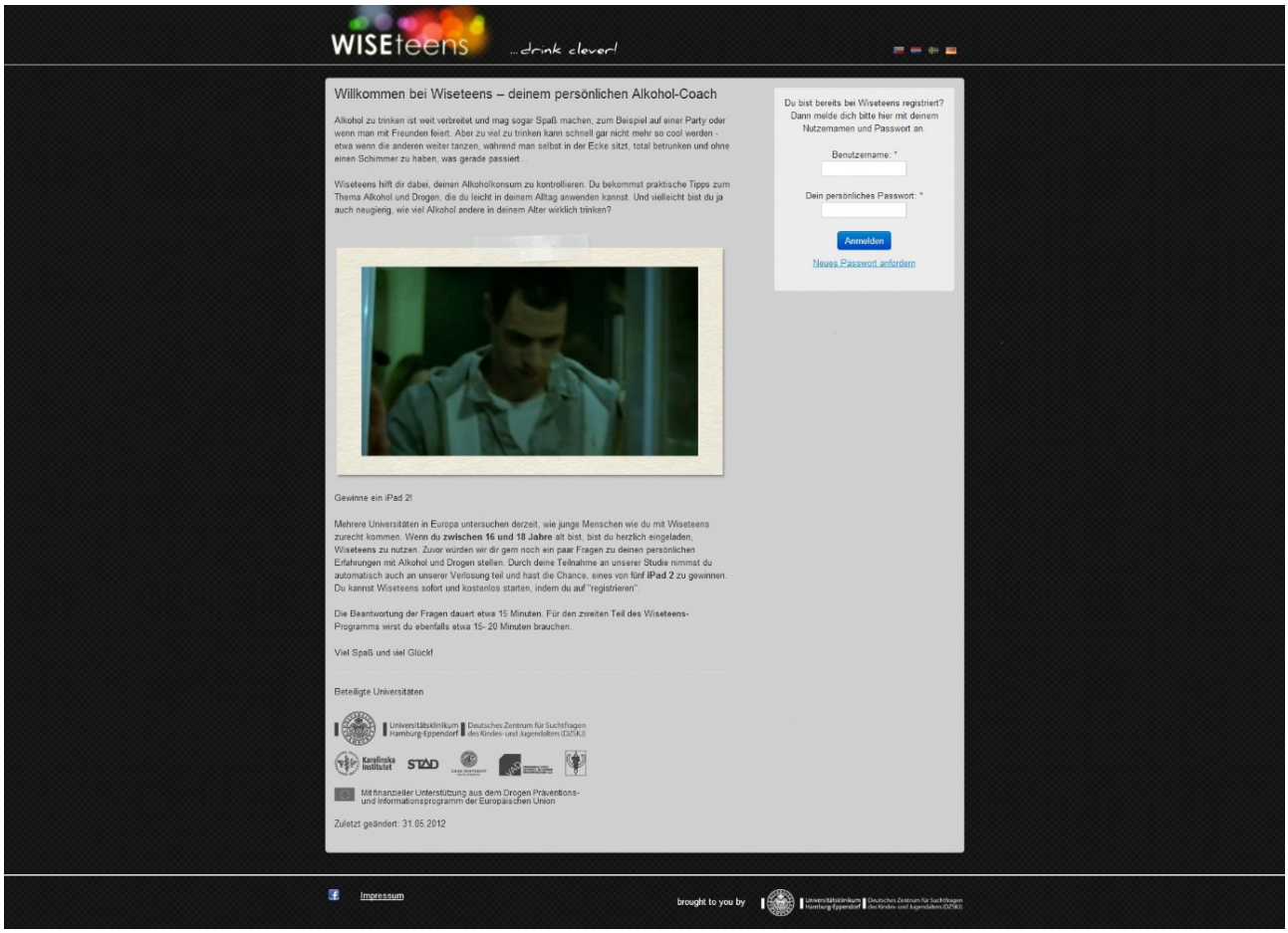


Figure 2. WISEteens landing page.



Recruitment

We promoted the open-access WISEteens landing page to recruit a convenience sample of potential participants using both online and offline strategies. As offline strategies, we developed print promotion materials (information leaflets and flyer cards) and distributed them in schools, youth-clubs, cafés, bars, stores, and adolescent specific events. We also used a three-fold online recruitment strategy with high rank of our websites' domain in widely used search engines, advertisements via popular social media, and links on affiliated health promotion sites. To motivate study participants and enhance follow-up rates, we promoted and held a prize draw for tablet computers among participants who provided follow-up assessment.

Procedure and Randomization

The participants were anonymous throughout the study. At the first visit, they were asked to register, which required a user name, email address, and a password that did not contain their name. On the landing page, they could choose their respective language flag for a different language than pre-defined by browser options. After registration, respondents were screened for at-risk substance use. Those fulfilling the inclusion criteria then received study information including confidentiality, voluntariness of participation, and data security, as well as information about the randomization protocol. Participants were not blinded to random allocation. After informed online consent, the baseline assessment was completed, including those items that form the base for tailoring intervention content. Participants in the intervention condition received a login code to enable exit and re-entrance. Randomization was generated automatically by an online computer program without stratification. The envisioned number of participants was sufficient to ensure randomization integrity and a likely balanced distribution among the two parallel groups [55].

Screening

An adapted version of the 6-item CRAFFT (Car, Relax, Alone, Forget, Friends, Trouble) tool was used to screen for at-risk use of alcohol and other drugs (see Multimedia Appendix 2). This tool has proven criterion validity compared to other screening tools [56] and is recommended for identification of at-risk adolescents [17]. A CRAFFT score with at least two positive items was the criterion for study inclusion [57].

Intervention

The WISEteens intervention relied on an interactive system to generate individually tailored content. All system-generated information was presented in small units that combined text and graphics (eg, photos and illustrative drawings) and directly referred to the participant's statements assessed in the first place (eg, substance use, sex, weight, perceptions of normative drinking). Navigation through the program was designed as a dialogue between the user and a virtual expert with "gates" (ie, choice options) at the end of each page to permit varying degrees of approval or disapproval with page content. The system used these responses to introduce subsequent content on the next page.

Intervention content, dialogue tone, and style was based on MI [40] and consisted of the following 6 components (see Multimedia Appendix 3 for screenshots), with the first three applying to alcohol but not illegal drug use and polydrug use: (1) feedback for individual drinking patterns with information on associated health and developmental risks, (2) normative feedback to descriptive drinking norms about sex- and age-matched peer drinking levels using graphed comparative information, (3) feedback for blood alcohol concentration (BAC) and associated health and other risks (ie, traffic crash, unintended sex) for the peak drinking episode in the last 2 weeks, (4) importance and confidence rulers with a short summary and feedback to encourage change readiness and exploration of personal strengths, resources, and volitional strategies for goal attainment, (5) decisional balance for selection of personal costs and benefits of current substance use and a subsequent graphical display of comparative gains and losses of behavior change in a balance sheet to illustrate ambivalence, and (6) identification and selection of personal high-risk situations for substance use and provision of behavioral strategies, for example, to resist peer pressure (the assumed mechanisms for change are displayed in Multimedia Appendix 4; for a more detailed description, see [54]).

The WISEteens intervention was pilot-tested in two steps. First, 10 adolescents chose their preferred design concept among three options. A preliminary version with the preferred "look & feel" was then pre-tested by 37 other adolescents to ensure ease of registration and navigation use, comprehensibility of intervention content, satisfaction with layout and design, appropriateness of dialogue style (eg, avoiding judgmental and confronting language), overall satisfaction with the program, and time to complete baseline assessment and intervention. Furthermore, open feedback, technical problems, translational ambiguities, and other problems were documented and the program was adapted accordingly. Median for assessment and intervention was 15 minutes, ranging from 5-30 minutes. Most adolescents were satisfied or totally satisfied with the system, design, comprehensibility, and intervention dialogue. Table 1 provides a summary of the pilot-results. The control group received assessment only and was directed to a short information page on where to find help in case of urgent counseling or medical needs.

Table 1. Results pilot-test (N=37[a]).

Items	Response					
	Mean (SD)	Yes, %	Neutral, %	No, %	Median	Range
Overall satisfaction with the program[b] (1="not satisfied at all", 4="totally satisfied")	3.1 (0.50)					
Acceptance of program layout and design[c] (1="not satisfied at all", 4="totally satisfied")	3.2 (0.66)					
Program comprehensibility[d] (1="not satisfied at all", 4="totally satisfied")	2.9 (0.80)					
Acceptance of dialogue style and tone						
"preachy"		0	43.8	56.3		
"non-judgmental"		68.8	12.5	18.8		
"appropriate"		81.3	12.5	6.3		
Duration to complete baseline assessment	16.5 min (5.5)				15 min	10-25 min
Duration to complete intervention	15.5 min (7.1)				15 min	5-30 min

[a]Mean age 16.38 (SD 1.23) years; 81.3% men.

[b]Single item on overall satisfaction.

[c]9 items (eg, login/registration procedure, clarity/structure, text amount, graphic design; Cronbach α=.89).

[d]7 items (eg, content comprehensibility, response format, information amount; Cronbach α=.77).

Data Collection and Measures

All study measures were administered anonymously and online via self-reports at baseline (t_0 before randomization) and at follow-up (t_1, 3 months after t_0) and required registration with a valid email address. Three months after completing the baseline assessment, participants were automatically invited to participate in the follow-up assessment and guided by an integrated hyperlink in the email invitation with one reminder email after 1 week.

Participants were asked for age, gender, country of residence (Sweden, Germany, Belgium, Czech Republic or other), current school attendance (yes/no), parental educational attainment (low, middle, high), their weight in kilograms for BAC-level feedback, descriptive norms [58, 59], and 12-months scores on the Alcohol Use Disorder Identification Test Consumption subscale (AUDIT-C [60]). To address baseline change motivation as a potential confounder, we included intention to reduce drinking (in the next 30 days) using a single item ("I intend to reduce my drinking during the next 30 days", 1="totally disagree" to 7="totally agree") and intention to

abstain from illegal drug use ("I intend to abstain from taking illegal drugs during the next 30 days", 1="totally disagree" to 4="totally agree") [57], as a continuous measure for change readiness [61].

Substance Use

All outcome measures concerned use in the past 30 days. Change in alcohol use (frequency, frequency of binge drinking, and quantity) between the two assessments was the primary outcome and measured based on the three items of the AUDIT-C screening tool [60]. This measure provides a widely used and valid index sum score for problem alcohol use of adolescents [62]. The three indicators are drinking frequency ("How often did you have a drink containing alcohol?"; 0="never" to 4="four or more times a week "), binge drinking frequency ("How often did you have 5 [4 for girls] or more drinks on one occasion, like during a party or on one night?"; 0="never" to 4="four or more times a week"), and drinking quantity ("How many drinks containing alcohol did you have on a typical day when you were drinking?"; 0="one or two" to 4="ten or more"). To assess the number of consumed drinks, we used a graphical display of various types of drinks with the indication to select the number of each drink per typical drinking occasion to account for national differences in typical standard drinks. Standard drinks were overall defined as containing 10-12 grams of pure ethanol, and responses were recoded to match the original 0-4 point scale. Using an index for drinking has two advantages over separate measures of alcohol use. First, it allows for modeling one drinking measure to increase the statistical power to detect intervention effects while maintaining several indicators of risky drinking. Second, due to the scale, characteristic outcome data can be analyzed as continuous data, which makes interpretation easier compared to discrete count drinking outcomes [63, 64]. We included the three separate outcomes (frequency, frequency of binge drinking, and quantity) and frequency of illegal drug use (1="never" to 5="four or more times a week" [65]) as secondary outcomes and past 30 days prevalence of illegal drug use and polydrug use as additional secondary outcomes. Because most drug use combinations in Europe include alcohol [10], we defined polydrug use as a dichotomous measure for use of alcohol and any illegal drug during the last 30 days, similar to previous studies [9].

Sample Size

Sample size calculation was based on the primary outcome with regard to effect sizes for alcohol use revealed by a recent review on Web-based interventions for young people [31]. According to results from similar studies, we expected a small effect size (Cohen's d=0.2). To reach power of 80% at a type I error rate of 5% in a two-sided test and expecting a dropout rate of approximately 50% [27, 66], we aimed at N=400 per intervention condition [67]. Possible country dependent clustering effects were not included in the sample size calculation because the study was designed as an individual-based RCT and higher level effects from four clusters were not previsioned [68]. Nevertheless, possible higher order country effects were addressed in all further analyses as described below.

Statistical Analyses

We first analyzed data on sample characteristics using t-tests (for metric data) and chi-square tests (for categorical data) to test for differences between intervention conditions. Next, we performed logistic regressions with completers (ie, those who provided valid follow-up data) versus dropouts as the binary dependent variable to test for possible attrition bias using all available sociodemographic and substance use variables as predictors. Intervention effects for primary and continuous secondary outcomes were tested using Linear Mixed Models (LMM) and binary secondary outcomes were analyzed using Generalized Linear Mixed Models. In all LMMs, we used change in outcome scores from baseline as the dependent variable, intervention condition as the only independent variable, baseline values as covariates ("fixed effects"), and country of residence as a single random effect ("random intercept"). This model controls for the correlation between baseline and follow-up outcome scores and does not require a repeated statement (ie, "time") and no time x group interaction term to interpret intervention effects. Because we found no higher order effect for country of residence (primary outcome: Wald Z=.77, P=.441), we skipped the random effect and adjusted the analyses for country of residence and possible confounders (ie, variables that were not balanced between intervention and control group at baseline) and predictors for missing data as additional covariates, which resulted in improved model fit (delta Bayesian information criterion=-13.16) (see [69] for a similar approach). Binary secondary outcome analyses (prevalence illegal drug use and polydrug use) focused on follow-up outcome values rather than change scores. All analyses are based on a complete-case dataset and an intention-to-treat (ITT) sample with imputation of missing follow-up data based on expectation maximization (EM). Both results are relevant and commonly reported in Web-based interventions particularly when dropout is large [70]. EM is a single imputation method that was shown to outperform the multiple imputation module available in SPSS in eHealth studies with high dropout rates [66]. Given the huge dropout in this study, we cross-checked the ITT-outcome analysis using a full information maximum likelihood (FIML) estimator for missing follow-up data provided in the structural equation model software AMOS to reduce estimate bias for missing data [71] and increase the robustness of findings (see Multimedia Appendix 5). For all analyses, we report estimated marginal means (percentages) and Cohen's d effect sizes. Distributions of outcomes (ie, skew and kurtosis) and missing-at-random requirements of missing data were checked prior to the main analyses. Results with a type I error rate of P<.05 in 2-sided tests were considered as statistically significant without adjustment for multiple comparisons but reporting of exact Pvalues [72]. All analyses other than in AMOS were performed using SPSS statistical software package version 22 [73].

Results

Sample Characteristics and Preliminary Analyses

The trial profile is shown in Figure 1. A total of 2673 participants logged on the WISEteens Web portal and participated in the initial screening. We excluded 655 (24.5%)

from the study due to a negative CRAFFT screening. This resulted in 2018 (75.5%) adolescents who gave consent to participate in the study and started subsequent baseline assessment (t_0). A total of 569 (28.1%) dropped out during the baseline assessment leaving 1449 participants who completed baseline assessment and were randomized to either the intervention (N=715) or control group (N=734). In the intervention group, 453 (63.4%) completed the brief intervention as measured by a log file record whether the last page of the intervention has been visible to the user. A total of 211 adolescents participated in the follow-up assessment after 3 months, corresponding to a valid response rate of 14.5%. In this subsample, the completion rate for the brief intervention was higher than in the full randomized sample (82.4%).

In the randomized sample (ITT population), the mean age was 16.8 years (SD 0.74), nearly half of the participants were women and the majority were currently attending school. Most participants were recruited in the Czech Republic due to a more intense offline recruitment in this country, indicating that adjustment of country of residence as an additional covariate was required in subsequent analyses. Participants in the intervention group tended to have a higher rate of parental educational attainment compared to the control group as indicated by a near significant difference (*P*=.060). This variable was therefore adjusted as an additional covariate in subsequent analyses [68]. With the exception of binge drinking frequency (*P*=.048), there were no significant group differences with regard to demographic or assessment data at baseline (see Table 2). Importantly, there were no baseline differences in the intention to change current alcohol and illegal drug use among the groups. Distributions of all continuous outcome variables showed acceptable skew and kurtosis with values well below 1.0 for baseline and follow-up assessments [74]. Regarding characteristics of participants who provided data at follow-up, group comparisons revealed no significant differences in any assessed variable (see Table 3). Overall baseline group comparisons thus indicate that the randomization was successful and that the completer-only subsample appears largely similar to the randomized sample.

Response rates were very similar for the intervention group (15%) and the control group (14%). Logistic regression analyses with attrition at follow-up as the dependent variable and all demographic (country, parental educational attainment, sex, age) and substance use related variables (all primary and secondary outcomes) as predictors explained 8.4% of the total response variance (Nagelkerke's R^2). Corresponding odds ratios (OR) revealed country of residence as the only significant predictors for dropout. Response rates were significantly lower for participants from the Czech Republic (11%) than for those from Sweden (23%, 58/251; OR 2.42, 95% CI 1.52-3.85, *P*<.001) and Germany (27%, 31/146; OR 2.514, 95% CI 1.41-4.49; *P*<.001) but not Belgium (18%, 26/143; OR 1.46, 95% CI 0.76-2.80; *P*=.26). The analysis thus indicates no biased attrition based on the available variables except for country of residence, which was adjusted in all subsequent analyses as a relevant predictor of missing data [75]. Analogous attrition analyses for intervention completion (as indicated by a record whether a user has "seen" the last page of the intervention dialogue) indicated that 12.2% of the total response variance was predicted by all study variables, with significantly more completers being women (*P*=.007) and significant differences between participants depending on country of residence (*P*<.001) with the highest intervention completion in Germany (89.7%), followed by Belgium (71.4%), Sweden (69.8), and the Czech Republic (56.0%).

Table 2. Baseline values for participant demographic and substance use related variables by intervention condition (randomized sample N=1449).

	Intervention	Control	*P* [a]
Randomized sample, n[b] (%)	715 (49.3)	734 (50.6)	
Age in years, mean (SD)	16.81 (0.75)	16.85 (0.74)	.253
Sex (women), %	47.8	48.6	.758
Country of residence, %			.574
Sweden	16.2	18.4	
Germany	10.9	9.3	
Belgium	9.8	9.9	
Czech Republic	63.3	62.4	
School status (yes), %	95.0	94.8	.885
Parental education level [c], %			.060
Low	10.1	10.9	
Middle	61.4	66.3	
High	28.5	22.9	
Intention to reduce alcohol[d], mean (SD)	3.04 (2.30)	3.22 (2.31)	.136
Intention to abstain from illegal drugs[d], mean (SD)	5.31 (2.46)	5.40 (2.40)	.501
Descriptive peer drinking norms, mean (SD)	2.31 (0.79)	2.32 (0.78)	.714
Substance use related risk (CRAFFT sum score), mean (SD)	2.75 (1.42)	2.72 (1.35)	.608
Age at first alcohol use, mean (SD)	12.92 (2.30)	13.00 (2.20)	.497
Alcohol use[e] (last 12 months), mean (SD)	4.91 (2.49)	5.10 (2.58)	.167
Alcohol use[f] (last 30 days), mean (SD)	5.43 (2.74)	5.46 (2.82)	.803
Drinking frequency[g], mean (SD)	2.01 (0.84)	1.97 (0.89)	.413
Drinking quantity[g], mean (SD)	1.79 (1.45)	1.84 (1.48)	.586
Binge drinking frequency[g], mean (SD)	1.67 (0.92)	1.78 (0.91)	.048
Illegal drug use (last 30 days), %	45.0	43.0	.460
Polydrug use (last 30 days), %	49.2	40.1	.734

[a]Results of chi-square tests for categorical and *t*tests for continuous measures.

[b]May differ for individual variables due to single missing values.

[c]Father's highest educational attainment.

[d]Scores ranging from 1-7 with higher scores indicating higher motivation for change.

[e]AUDIT-C index score, past 12 months.

[f]AUDIT-C based index score, past 30 days (primary outcome).

[g]Separate drinking indicators, scores ranging from 0-4 with higher scores indicating more severe drinking.

Table 3. Baseline values for participant demographic and substance use related variables by intervention condition (completers-only sample N=211).

	Intervention	Control	*P* [a]
Completers-only sample, n[b] (%)	108 (51.2)	103 (48.8)	
Age in years, mean (SD)	16.87 (0.71)	17.03 (0.76)	.130
Sex (women), %	52.9	52.6	.959
Country of residence, %			.955
Sweden	26.9	28.2	
Germany	15.7	13.6	
Belgium	13.0	11.7	
Czech Republic	44.4	46.6	
School status (yes)	88.0	88.3	.961
Parental education level [c], %			.198
Low	14.6	12.4	
Middle	43.8	57.3	
High	41.6	30.3	
Intention to reduce alcohol[d], mean (SD)	2.76 (2.12)	3.13 (2.08)	.223
Intention to abstain from illegal drugs[d], mean (SD)	5.39 (2.37)	5.59 (2.22)	.539
Descriptive peer drinking norms, mean (SD)	2.33 (0.66)	2.35 (0.64)	.856
Substance use related risk (CRAFFT sumscore), mean (SD)	2.56 (1.32)	2.80 (1.35)	.210
Age at first alcohol use, mean (SD)	12.97 (2.34)	13.06 (2.27)	.781
Alcohol use[e] (last 12 months), mean (SD)	4.89 (2.27)	5.20 (2.47)	.348
Alcohol use[f] (last 30 days), mean (SD)	5.35 (2.44)	5.53 (2.87)	.644
Drinking frequency[g], mean (SD)	2.09 (0.79)	2.04 (0.89)	.695
Drinking quantity[g], mean (SD)	1.77 (1.27)	1.85 (1.46)	.694
Binge drinking frequency[g], mean (SD)	1.54 (0.86)	1.69 (0.93)	.261
Illegal drug use (last 30 days), %	42.6	39.5	.946
Polydrug use (last 30 days), %	37.0	31.1	.313

[a]Results of chi-square tests for categorical and *t*tests for continuous measures.

[b]May differ for individual variables due to single missing values.

[c]Father's highest educational attainment.

[d]Scores ranging from 1-7 with higher scores indicating higher motivation for change.

[e]AUDIT-C index score, past 12 months.

[f]AUDIT-C based drinking index score, past 30 days (primary outcome).

[g]Separate drinking indicators, scores ranging from 0-4 with higher scores indicating more severe drinking.

Primary Outcome: Past Month Drinking Index

Tables 4 and 5 report the primary and secondary intervention outcomes of this trial at follow-up based on the non-imputed completer-sample and Tables 6 and 7 for the EM-imputed intention-to-treat-sample. All analyses concern substance use in the past 30 days and were adjusted for baseline scores, country of residence, and parental educational attainment.

Based on the non-imputed sample 3 months after the intervention, participants in the intervention group have reduced their drinking as indicated by reduced AUDIT-C based scores relative to baseline with an adjusted mean change of -0.85 (95% CI -1.49 to -0.26) while those in the control group slightly increased their drinking as indicated by a mean increase in drinking of 0.16 (95% CI -0.50 to 0.82). Adjusted mean differences between both groups were 1.02 (95% CI 0.25-1.79) and statistically significant ($F_{1,134}$=6.8, P=.010, d=.26). The corresponding between-group effect was smaller in the (imputed) ITT analysis due to significant reductions relative to baseline in the control group. However, the significant between-group effect (0.16, 95% CI 0.02-0.25) was maintained ($F_{1,1329}$=5.2, P=.022, d=.04). Additional analysis in AMOS based on an FIML estimation for missing outcome assessments

confirmed these results (B=-0.72, β=-0.13, *P*=.046; see Multimedia Appendix 5).

Secondary Outcomes: Drinking (AUDIT-C Based Separate Items)

We conducted identical analyses for the three drinking indicators separately. In the non-imputed sample, we found a significant between-group difference in drinking frequency of 0.25 (95% CI 0.02-0.50) in favor of the WISEteens group ($F_{1,134}$=4.4, *P*=.037, *d*=.15), which was not maintained in the analysis based on EM-imputation ($F_{1,1329}$=3.2, *P*=.073, *d*=.11) due to significant reductions relative to baseline in both groups (*P*s<.001). We obtained a similar result for binge drinking frequency with a significant adjusted mean difference between groups of 0.31 (95% CI 0.01-0.61; $F_{1,121}$=4.2, *P*=.044, *d*=.16) in the non-imputed data analysis, which was not maintained in the imputed analysis ($F_{1,1329}$=2.3, *P*=130, *d*=.01). Additional FIML analysis for both outcomes revealed non-significant and similar results as in the imputed analysis (drinking frequency: B=-0.13, β=-0.08, *P*=.230 and binge drinking frequency: B=-0.25, β=-0.14, *P*=.059; see Multimedia Appendix 5). For drinking quantity, we found significant reductions relative to baseline for the intervention group in the non-imputed analysis of -0.39 (95% CI -0.72 to -0.06, *P*=.024), but these reductions were non-significantly different from the control group (between-group differences: 0.31, 95% CI -0.17 to 0.62; $F_{1,155}$=1.3, *P*=.257, *d*=.13). This effect was significant in the EM-imputed data-set ($F_{1,1329}$=5.3, *P*=.021, *d*=.05). However, when cross-checked using the FIML approach employed in AMOS, these differences were no longer significant (B=-0.22, β=-0.08, *P*=.209, see Multimedia Appendix 5).

Secondary Outcomes: Illegal Drug Use and Polydrug Use

Results for frequency and prevalence of illegal drug use as well as polydrug use prevalence are summarized in Tables 5 and 7. With regard to these outcomes, we found no statistically significant between-group effects in the non-imputed (*P*s=.138-.311) and the EM-imputed (*P*s=.363-.871) datasets. Although overall, both groups show numerical decreases between the measurements that were statistically significant in the intervention group for illegal drug use prevalence (*P*=.025) and polydrug use prevalence (*P*=.012) in the non-imputed analyses and statistically significant for all 3 outcomes (ie, frequency of illegal drug use and illegal drug use and polydrug prevalence, *P*s<.001) in the imputed analyses.

Table 4. Intervention effects[a] on primary and (continuous) secondary outcomes (non-imputed dataset[b]).

Outcomes after 3 months	WISEteens group (n=715)			Control group (n=734)			Between-group differences			
	Mean (SD)	Change from baseline, adjusted mean (95% CI)	*P*	Mean (SD)	Change from baseline, adjusted mean (95% CI)	*P*	Adjusted mean (95% CI)	*F*(df)	*P*	*d*
Index alcohol use [c]										
Baseline	5.43 (2.74)			5.46 (2.81)						
3-months follow-up	4.59 (2.77)	-0.85 (-1.49 to -0.26)	.009	5.35 (2.57)	0.16 (-0.50 to 0.82)	.614	1.02 (0.25 to 1.79)	6.80 (1, 134)	.010	.26
Drinking frequency [d]										
Baseline	2.01 (0.84)			1.97 (.90)						
3-months follow-up	1.80 (0.84)	-0.36 (-0.55 to -0.16)	<.001	1.88 (0.81)	-0.11 (-0.31 to 0.10)	.305	0.25 (0.02 to 0.50)	4.40 (1, 144)	.037	.15
Binge drinking frequency [d]										
Baseline	1.67 (0.92)			1.78 (0.91)						
3-months follow-up	1.39 (0.95)	-0.11 (-0.36 to 0.14)	.375	1.66 (0.85)	0.20 (-0.07 to 0.47)	.152	0.31 (0.01 to 0.61)	4.20 (1, 121)	.044	.16
Drinking quantity [d]										
Baseline	1.79 (1.45)			1.84 (1.48)						
3-months follow-up	1.59 (1.39)	-0.39 (-0.72 to -0.06)	.024	1.83 (1.36)	-0.16 (-0.50 to 0.17)	.336	0.23 (-0.17 to 0.62)	1.30 (1, 155)	.257	.13
Illegal drug use frequency [e]										
Baseline	0.87 (1.20)			0.80 (1.14)						
3-months follow-up	0.69 (1.10)	-0.13 (-0.37 to 0.11)	.292	0.71 (1.07)	-0.03 (-0.28 to 0.22)	.805	0.10 (-0.21 to 0.40)	0.40 (1, 133)	.532	.07

[a]Based on linear mixed model with group as fixed factor, changes from baseline as outcomes, and baseline scores, country, and parental educational attainment as covariates for continuous outcomes. Cohen's *d* calculated by subtracting the average difference score between pretest and posttest of the control group from the corresponding difference score of the intervention group, and dividing the result by the pooled standard deviation of the baseline scores.

[b]Valid follow-up data for n=211 trial participants.

[c]Adapted AUDIT-C index score (primary outcome).

[d]Adapted AUDIT-C indicators, scores ranging from 0-4 with higher scores indicating more severe drinking.

[e]Scores ranging from 0-4 with higher scores indicating more frequent illegal drug use.

Table 5. Intervention effects[a] on binary secondary outcomes (non-imputed dataset[b]).

Outcome after 3 months	WISEteens group (n=715)		Control group (n=734)		Between-group differences		
	% (SE)	*P*	% (SE)	*P*	*F*(df)	*P*	OR (95% CI)
Illegal drug use prevalence (%)							
Baseline	45.0 (0.02)		43.0 (0.02)				
3-months follow-up	36.1 (0.05)	.025	39.5 (0.05)	.431	1.03 (1, 133)	.311	0.67 (0.31 to 1.45)
Polydrug [c] prevalence (%)							
Baseline	42.9 (0.02)		40.1 (0.02)				
3-months follow-up	31.3 (0.05)	.012	36.8 (0.05)	.235	2.22 (1, 163)	.138	0.57 (0.27 to 1.20)

[a]Based on (logistic) general linear mixed model with group as fixed factor, follow-up values as outcomes, and baseline scores, country, and parental educational attainment as covariates.

[b]Valid follow-up data for n=211 trial participants.

[c]Combined use of alcohol and any illegal drug in past 30 days.

Table 6. Intervention effects on primary and (continuous) secondary outcomes (EM-imputed dataset).

Outcomes after 3 months	WISEteens group (n=715)			Control group (n=734)			Between-group differences			
	Mean (SD)	Change from baseline, adjusted mean (95% CI)	*P*	Mean (SD)	Change from baseline, adjusted mean (95% CI)	*P*	Adjusted mean (95% CI)	*F*(df)	*P*	*d*
Index alcohol use [b]										
Baseline	5.24 (2.71)			5.25 (2.78)						
3-months follow-up	4.72 (1.58)	-0.63 (-0.73 to -0.52)	<.001	4.82 (1.52)	-0.49 (-0.60 to -0.39)	<.001	0.13 (0.02 to 0.25)	5.23 (1, 1329)	.022	.04
Drinking frequency [c]										
Baseline	1.98 (0.81)			1.93 (0.90)						
3-months follow-up	1.75 (0.47)	-0.24 (-0.27 to -0.20)	<.001	1.76 (0.46)	-0.20 (-0.24 to -0.17)	<.001	0.03 (-0.003 to -0.07)	3.21 (1, 1329)	.073	.11
Binge drinking frequency [c]										
Baseline	1.54 (0.99)			1.58 (1.02)						
3-months follow-up	1.39 (0.50)	-0.20 (-0.24 to -0.16)	<.001	1.42 (0.47)	-0.16 (-0.20 to -0.12)	<.001	0.03 (-0.01 to -0.08)	2.30 (1, 1329)	.130	.01
Drinking quantity [c]										
Baseline	1.74 (1.46)			1.75 (1.49)						
3-months follow-up	1.64 (0.77)	-0.15 (-0.20 to -0.10)	<.001	1.71 (0.77)	-0.08 (-0.14 to -0.03)	.001	0.07 (0.01 to 0.12)	5.33 (1, 1329)	.021	.05
Illegal drug use frequency [d]										
Baseline	0.84 (1.15)			0.76 (1.08)						
3-months follow-up	0.70 (0.76)	-0.12 (-0.15 to -0.08)	<.001	0.67 (0.71)	-0.11 (-0.14 to -0.08)	<.001	0.01 (-0.3 to 0.04)	0.18 (1, 1329)	.670	.04

[a]Based on linear mixed model with group as fixed factor, changes from baseline as outcomes, and baseline scores, country, and parental educational attainment as covariates for continuous outcomes. Cohen's d calculated by subtracting the average difference score between pretest and posttest of the control group from the corresponding difference score of the intervention group and dividing the result by the pooled standard deviation of the baseline scores.

[b]Adapted AUDIT-C index score (primary outcome).

[c]Adapted AUDIT-C indicators, scores ranging from 0-4 with higher scores indicating more severe drinking.

[d]Scores ranging from 0-4 with higher scores indicating more frequent illegal drug use.

Table 7. Intervention effects[a] on binary secondary outcomes (EM-imputed dataset[b]).

Outcomes after 3-months	WISEteens group (n=715)		Control group (n=734)		Between-group differences		
	% (SE)	*P*	% (SE)	*P*	*F*(df)	*P*	OR (95% CI)
Illegal drug usep revalence (%)							
Baseline	49.8 (0.02)		49.6 (0.02)				
3-months follow-up	41.7 (0.02)	<.001	39.8 (0.02)	<.001	1.30 (1, 1446)	.254	1.22 (0.87 to 1.73)
Polydrug [c] prevalence (%)							
Baseline	47.8 (0.02)		46.3 (0.02)				
3-months follow-up	41.1 (0.02)	<.001	39.8 (0.02)	<.001	0.02 (1, 1446)	.888	1.03 (0.73 to 1.44)

[a]Based on (logistic) general linear mixed model with group as fixed factor, follow-up values as outcomes, and baseline scores, country and parental educational attainment as covariates.

[b]The binary imputed prevalence outcomes (illegal drug use and polydrug use), a real number between 0 and 1 was transformed back into a dichotomous variable by rounding off to two positions behind the decimal point.

[c]Combined use of alcohol and any illegal drug in past 30 days.

Discussion

Principal Findings

The purpose of this study was to test the effectiveness of a fully automated Web-based screening and brief motivational intervention targeting adolescents with at-risk substance use in Europe.

We found that self-reported risk drinking as measured by a drinking index (ie, drinking frequency, frequency of binge drinking, and typical quantity of drinks) was significantly reduced for participants in the intervention group. The effect on the primary alcohol use outcome was consistent across imputation and non-imputation ("completers") based analyses but accented in the non-imputed data analysis, even though statistical power was low for 3-month effects due to large loss to follow-up assessment. Secondary analyses using the three drinking indicators as separate outcomes revealed statistically significant mean differences at follow-up in favor of the WISEteens intervention group for drinking frequency and binge drinking frequency but not quantity when missing follow-up data was not imputed. In contrast, analyses using an EM-imputed dataset revealed drinking quantity as the only significant secondary effect. For illegal drug use or polydrug use, there were no significant intervention effects.

The effect sizes obtained in this study are small but match those summarized in recent systematic reviews for fully automated interventions for young adults [32] and meta-analyses for single session interventions [76, 77]. Moreover, they correspond with effect sizes reported for face-to-face brief interventions for youth who use alcohol and other drugs [17, 78, 79] and indicate that expected effects of MI-based interventions on substance use may indeed be small [41] but can be relevant when a large population can be reached. Overall, there are currently few Web-based interventions targeting adolescents, which limits direct comparisons to prior studies. However, our study contradicts results from one recent RCT that tested the effects of a similar intervention (What Do You Drink [WDYD]). This trial targeted drinking among young people (15-20 years) with low educational background in the Netherlands [80]. While WDYD and WISEteens are comparable in central characteristic (eg, age group, cultural context, single session fully automated delivery mode, intervention duration, applied theory, and outcome measures), there are a number of differences that could account for the divergent effects, such as the school-based study implementation and different follow-up times. Moreover, in the WDYD trial participants with an indication of severe problem drinking at baseline were excluded, while in our study about half of the participants were above the AUDIT-C risk cut-off of 5 points [62] at baseline. Severity of baseline drinking can influence effects of brief motivational interventions with stronger effects among subgroups of heavy drinkers [81].

WISEteens was not effective to address illegal drug use and polydrug use adequately. However, the number of participants with drug and polydrug use was rather low in our sample and meaningful between-group effects might have been undetected due to insufficient statistical power [37]. Note that there were notable decreases in the prevalence rates for polydrug use and illegal drug use in the intervention group, while there were no changes and even slight increases in the control group. There were also no spill-over effects from reduced problem drinking to other substance use, which suggests that effects on targeted outcomes may not translate to untargeted outcomes [82, 83].

In general, effects from comparable Web-based interventions for illegal drugs are typically smaller than for alcohol [32]. Moreover, the main hypothesis and focus of the intervention was on drinking. Some intervention elements, such as decisional balance, importance, and confidence ruler and advice for risk situations were available for alcohol per default and participants with other drug use were explicitly encouraged to take the exercises as templates for use of other substances. Thus, the lack of positive outcomes may also be the result of limited specific intervention content for drugs other than alcohol, which may be a limitation. This notwithstanding, the lack of significant effects in our study corresponds to previous trials among students that failed to promote positive behavior change [84, 85], although such interventions can in principle be effective in the general population [32, 86].

Strengths and Limitations

Our study is among the first to report on a targeted fully automated Web-based brief intervention among at-risk adolescents with excessive alcohol drinking and drug use in a randomized controlled trial. From a public health perspective, the significant effect on drinking is relevant. Notwithstanding the often dramatic consequences associated with illegal drug use, alcohol is the most frequently used psycho-active substance during adolescence, alcohol use disorders are among the most prevalent and costly mental disorders in industrialized countries, and prevention is a public health priority [87]. Although the main burden of alcohol-related diseases and injuries becomes apparent in adulthood, it is well documented that early at-risk alcohol use can lead to persistent problems [88,89]. Considering the magnitude of youth with at-risk alcohol use the need for effective targeted prevention may be particularly high in Europe. For example in Germany, hospital admissions due to acute alcohol intoxication have increased substantially in the past years, although the proportion of youth who drink alcohol is decreasing [7]. With this study, we provide initial support for the effectiveness of Web-based brief interventions for adolescents in Europe—an approach that has proven useful in other critical heavy drinking populations, such as college students [31, 33].

The WISEteens intervention was designed to reflect valid face-to-face motivational strategies in a Web-based format. We extended individual feedback techniques typically used in Web-based brief interventions by other MI techniques from face-to-face interventions, such as decisional balance, confidence and importance ruler, and provision of behavioral and regulatory strategies to resist peer pressure [38, 40, 90]. Moreover, we aimed to mimic MI principles by applying a carefully designed and pilot-tested motivational dialogue to reflect possibly “empathic” language, acknowledgement of ambivalent goals, needs for autonomy and self-directedness, and appreciation of change but not necessarily abstinence. Although our results cannot address this issue directly, we assume that this may have contributed to the favorable effects revealed in this study. Certainly we acknowledge the limits in delivering MI-consistent methods such as relational factors in a fully automated format [91]. However, arguably the ability of such interventions to deliver standardized intervention content on a large scale and at low cost may outweigh these limitations.

Finally, even though response to follow-up assessment was very low, screening and intervention were accessed and completed by a reasonably large number of teenagers, which may indicate acceptance and ecological validity in the target group. The intervention completion rate indicates acceptable user engagement in the “real world” [92] and a good balance between the required amount of program exposure and adherence requirements of a single session Web-based delivery format for adolescents [50, 93].

Our study has a number of limitations. First and foremost, results are limited by the higher than expected dropout rate for follow-up assessment, which is a frequent problem in Web-based trials [66, 70]. Dropout might be partly caused by invalid email addresses used by the participants and the fact that the system sent only one email reminder per participant [94]. Even though we detected no serious attrition bias, this may limit the validity of the study findings. Although in case of large dropout, any approach to missing data imputation such as the EM method employed in this study could be compromised. We approached this problem by cross-check analyses using an FIML approach to missing data estimation, which yielded similar and significant results for the primary outcome. Moreover, it must be noted that attrition was mainly a problem for the evaluation (ie, attrition took place between measurement points) and much less for intervention adherence (ie, intervention completion rates). The fact that we were able to identify intervention effects, even with the fewer than required number of participants needed to detect small effects in the evaluation, could be interpreted as an underestimation of intervention effects. This notwithstanding and although the study was performed in four European countries, the results should be taken only under careful consideration and follow-up should be translated to other countries or even to other regions or to other groups of adolescents in the four involved countries. The inconsistent results of the imputed and non-imputed analyses and substantial between-analyses deviations in obtained effect sizes serve as a quantitative indicator of uncertainty in these results due to the substantial amount of missing follow-up data.

Another limitation is reliance on self-reported data in this study, which are often associated with underreporting of alcohol drinking and other drug use [95]. Measures to avoid underreporting were assurance of and advice for confidentiality and a non-judgmental and non-confronting MI style employed throughout the intervention. Moreover in our study, the self-report was given anonymously and without personal contact, which may add to the reliability of self-reported data. Moreover, we relied on a primary outcome that was based on the AUDIT-C, which was developed as a screening tool for harmful alcohol use in adults. Given the adapted reference time of drinking in the past 30 days, it is difficult to interpret the practical significance of the measured changes in outcome scores in response to the intervention. Furthermore, participants were not blinded to the assigned interventions, which is a common limitation in Web-based trials [96].

Finally, given the focus on intervention effects in this study, we have not systematically included recruitment and reach in our analyses although these are important issues for estimating public health impact [97]. Rather, we aimed for a convenience sample from the general population by employing a pragmatic recruitment procedure. However, the incitement by lottery as an incentive for participation may have increased the reach to a higher level than can be expected in implementation outside a research project. While the self-selected nature of our sample again limits generalization, we consider the realistic setting of this trial a significant strength. In fact, apart from the evaluation requirements at baseline, the actual intervention program was equivalent to a potential real world application. We thus feel confident in stating that our study has realistic public health implications.

Conclusion

Web-based interventions to reduce adolescent at-risk substance use hold promise for accessibility and large scale dissemination but have rarely been tested in randomized controlled trials. Although limited by substantial dropout to follow-up assessment, our findings imply that young adolescents with excessive drinking can benefit from targeted interventions based on MI techniques and counseling style in a fully automated Web-based delivery mode. However, we found no effect on drug use, which calls for further research on effective intervention models, delivery modes, and recruitment strategies.

Acknowledgments

The authors would like to especially thank Sonja Bröning, Magdalena Drechsel, Helena Hansson, and Ulla Zetterlind for their contribution to development of the intervention program; Petr Winkler and Nicklas Kartengren for assistance in data collection; Anne Berman and Veerle Soyez for project advisory; Hamoud Gahl for technical implementation; and Anne Daubmann and Levente Kriston for advice on data analysis.

The study was funded by the Drug Prevention and Information Programme of the European Union (Grant agreement No. JUST/2010/DPIP/AG/0914-30-CE-0379823/00-48).

Conflicts of Interest

The authors were involved in the development of the intervention.

Multimedia Appendix 1

CONSORT-eHEALTH V 1.6.1.

[PDF File (Adobe PDF File), 1MB - jmir_v18i5e103_app1.pdf]

Multimedia Appendix 2

CRAFFT Screening for at-risk substance use (A CRAFFT score of more than one positive item was inclusion criterion).

[PDF File (Adobe PDF File), 411KB - jmir_v18i5e103_app2.pdf]

Multimedia Appendix 3

Screenshots.

[PDF File (Adobe PDF File), 1MB - jmir_v18i5e103_app3.pdf]

Multimedia Appendix 4

Assumed behavior change model and targeted WISEteens content.

[PDF File (Adobe PDF File), 41KB - jmir_v18i5e103_app4.pdf]

Multimedia Appendix 5

Intervention effects at 3-month follow-up, ITT-analysis based on FIML.

[PDF File (Adobe PDF File), 47KB - jmir_v18i5e103_app5.pdf]

References

1. Bekkering GE, Aertgeerts B, Asueta-Lorente J, Autrique M, Goossens M, Smets K, et al. Practitioner review: evidence-based practice guidelines on alcohol and drug misuse among adolescents: a systematic review. J Child Psychol Psychiatry 2014 Jan;55(1):3-21. [doi: 10.1111/jcpp.12145] [Medline: 24117606]
2. Lim SS, Vos T, Flaxman AD, Danaei G, Shibuya K, Adair-Rohani H, et al. A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet 2012 Dec 15;380(9859):2224-2260 [FREE Full text] [doi: 10.1016/S0140-6736(12)61766-8] [Medline: 23245609]
3. Masten AS, Faden VB, Zucker RA, Spear LP. Underage drinking: a developmental framework. Pediatrics 2008 Apr;121 Suppl 4:S235-S251. [doi: 10.1542/peds.2007-2243A] [Medline: 18381492]
4. Thomasius R, Stolle M, Sack P. Entwicklungspsychopathologisches Modell. In: Thomasius R, Schulte-Markwort M, Küstner UJ, Riedesser P, editors. Handbuch der Suchtstörungen im Kindes- und Jugendalter. Stuttgart: Schattauer; 2009:139-146.

5. Heron J, Macleod J, Munafò MR, Melotti R, Lewis G, Tilling K, et al. Patterns of alcohol use in early adolescence predict problem use at age 16. Alcohol Alcohol 2012;47(2):169-177 [FREE Full text] [doi: 10.1093/alcalc/agr156] [Medline: 22215001]
6. Wittchen H, Behrendt S, Höfler M, Perkonigg A, Lieb R, Bühringer G, et al. What are the high risk periods for incident substance use and transitions to abuse and dependence? Implications for early intervention and prevention. Int J Methods Psychiatr Res 2008 Jun;17 Suppl 1:S16-S29. [doi: 10.1002/mpr.254] [Medline: 18543359]
7. Diestelkamp S, Arnaud N, Sack PM, Wartberg L, Daubmann A, Thomasius R. Brief motivational intervention for adolescents treated in emergency departments for acute alcohol intoxication - a randomized-controlled trial. BMC Emerg Med 2014;14:13 [FREE Full text] [doi: 10.1186/1471-227X-14-13] [Medline: 24975110]
8. Hibell B, Guttormsson U, Ahlström S, Balakireva O, Bjarnason T, Kokkevi A. ESPAD report. Stockholm: The Swedish council for information on alcohol and other drugs (CAN); 2012. The 2011 ESPAD report: Substance use among students in 36 European countries URL: http://www.espad.org/uploads/espad_reports/2011/the_2011_espad_report_full_2012_10_29.pdf[accessed [accessed 2016-04-18] [WebCite Cache ID 6grRpZ7UG]
9. Olszewski D, Matias J, Monshouwer K, Kokkevi A. Polydrug use among 15- to 16-year olds: Similarities and differences in Europe. Drugs: Education, Prevention and Policy 2010 Jun 30;17(4):287-302. [doi: 10.3109/09687630902806715]
10. EMCDDA. European Drug Report. Lisbon: European Monitoring Centre for Drugs and Drug Addiction (EMCDDA); 2009. Annual Report 2009: The State of the Drug Problem URL: http://www.emcdda.europa.eu/publications/annual-report/2009[accessed [accessed 2016-04-18] [WebCite Cache ID 6grSJFM3H]
11. Viner RM, Taylor B. Adult outcomes of binge drinking in adolescence: findings from a UK national birth cohort. J Epidemiol Community Health 2007 Oct;61(10):902-907 [FREE Full text] [doi: 10.1136/jech.2005.038117] [Medline: 17873228]
12. Van Der Vorst H, Vermulst AA, Meeus WHJ, Deković M, Engels RCME. Identification and prediction of drinking trajectories in early and mid-adolescence. J Clin Child Adolesc Psychol 2009 May;38(3):329-341. [doi: 10.1080/15374410902851648] [Medline: 19437294]
13. Pharo H, Sim C, Graham M, Gross J, Hayne H. Risky business: executive function, personality, and reckless behavior during adolescence and emerging adulthood. Behav Neurosci 2011 Dec;125(6):970-978. [doi: 10.1037/a0025768] [Medline: 22004262]
14. Miller JW, Naimi TS, Brewer RD, Jones SE. Binge drinking and associated health risk behaviors among high school students. Pediatrics 2007 Jan;119(1):76-85. [doi: 10.1542/peds.2006-1517] [Medline: 17200273]
15. Hale DR, Fitzgerald-Yau N, Viner RM. A systematic review of effective interventions for reducing multiple health risk behaviors in adolescence. Am J Public Health 2014 May;104(5):e19-e41. [doi: 10.2105/AJPH.2014.301874] [Medline: 24625172]
16. Bühler A, Thrul J. Expertise zur Suchtprävention - Aktualisierte und erweiterte Neuauflage der Expertise zur Prävention des Substanzmissbrauchs. Köln: BZgA; 2013.
17. Patton R, Deluca P, Kaner E, Newbury-Birch D, Phillips T, Drummond C. Alcohol screening and brief intervention for adolescents: the how, what and where of reducing alcohol consumption and related harm among young people. Alcohol Alcohol 2014;49(2):207-212 [FREE Full text] [doi: 10.1093/alcalc/agt165] [Medline: 24232178]
18. Meier MH, Caspi A, Houts R, Slutske WS, Harrington H, Jackson KM, et al. Prospective developmental subtypes of alcohol dependence from age 18 to 32 years: implications for nosology, etiology, and intervention. Dev Psychopathol 2013 Aug;25(3):785-800 [FREE Full text] [doi: 10.1017/S0954579413000175] [Medline: 23880392]
19. Van HS, Harris SK, Brooks T, Carey P, Kossack R, Kulig J, New England Partnership for Substance Abuse Research. The "Six T's": barriers to screening teens for substance abuse in primary care. J Adolesc Health 2007 May;40(5):456-461. [doi: 10.1016/j.jadohealth.2006.12.007] [Medline: 17448404]
20. Rumpf H. Frühintervention bei problematischem Alkoholkonsum in der medizinischen Versorgung: Auf dem Weg zur Implementierung. SUCHT 2009 Jan 21;55(6):326-327. [doi: 10.1024/2009.06.01]
21. Cunningham J, Breslin F. Only one in three people with alcohol abuse or dependence ever seek treatment. Addict Behav 2004 Jan;29(1):221-223. [Medline: 14667433]
22. Flory K, Lynam D, Milich R, Leukefeld C, Clayton R. Early adolescent through young adult alcohol and marijuana use trajectories: early predictors, young adult outcomes, and predictive utility. Dev Psychopathol 2004;16(1):193-213. [Medline: 15115071]
23. Murray E. Web-based interventions for behavior change and self-management: potential, pitfalls, and progress. Med 2 0 2012;1(2):e3 [FREE Full text] [doi: 10.2196/med20.1741] [Medline: 25075231]
24. Riper H, Andersson G, Christensen H, Cuijpers P, Lange A, Eysenbach G. Theme issue on e-mental health: a growing field in internet research. J Med Internet Res 2010;12(5):e74 [FREE Full text] [doi: 10.2196/jmir.1713] [Medline: 21169177]
25. Ritterband LM, Tate DF. The science of internet interventions. Introduction. Ann Behav Med 2009 Aug;38(1):1-3. [doi: 10.1007/s12160-009-9132-5] [Medline: 19816750]
26. Burns JM, Davenport TA, Durkin LA, Luscombe GM, Hickie IB. The internet as a setting for mental health service utilisation by young people. Med J Aust 2010 Jun 7;192(11 Suppl):S22-S26. [Medline: 20528703]

27. Spijkerman R, Roek MAE, Vermulst A, Lemmers L, Huiberts A, Engels RCME. Effectiveness of a web-based brief alcohol intervention and added value of normative feedback in reducing underage drinking: a randomized controlled trial. J Med Internet Res 2010;12(5):e65 [FREE Full text] [doi: 10.2196/jmir.1465] [Medline: 21169172]
28. 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]
29. White A, Kavanagh D, Stallman H, Klein B, Kay-Lambkin F, Proudfoot J, et al. Online alcohol interventions: a systematic review. J Med Internet Res 2010;12(5):e62 [FREE Full text] [doi: 10.2196/jmir.1479] [Medline: 21169175]
30. Bewick BM, Trusler K, Barkham M, Hill AJ, Cahill J, Mulhern B. The effectiveness of web-based interventions designed to decrease alcohol consumption--a systematic review. Prev Med 2008 Jul;47(1):17-26. [doi: 10.1016/j.ypmed.2008.01.005] [Medline: 18302970]
31. Tait RJ, Christensen H. Internet-based interventions for young people with problematic substance use: a systematic review. Med J Aust 2010 Jun 7;192(11 Suppl):S15-S21. [Medline: 20528701]
32. Tait RJ, Spijkerman R, Riper H. Internet and computer based interventions for cannabis use: a meta-analysis. Drug Alcohol Depend 2013 Dec 1;133(2):295-304. [doi: 10.1016/j.drugalcdep.2013.05.012] [Medline: 23747236]
33. Riper H, Spek V, Boon B, Conijn B, Kramer J, Martin-Abello K, et al. Effectiveness of E-self-help interventions for curbing adult problem drinking: a meta-analysis. J Med Internet Res 2011;13(2):e42 [FREE Full text] [doi: 10.2196/jmir.1691] [Medline: 21719411]
34. 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]
35. Carey KB, Scott-Sheldon LAJ, Carey MP, DeMartini KS. Individual-level interventions to reduce college student drinking: a meta-analytic review. Addict Behav 2007 Nov;32(11):2469-2494 [FREE Full text] [doi: 10.1016/j.addbeh.2007.05.004] [Medline: 17590277]
36. Fairlie AM, Wood MD, Laird RD. Prospective protective effect of parents on peer influences and college alcohol involvement. Psychol Addict Behav 2012 Mar;26(1):30-41 [FREE Full text] [doi: 10.1037/a0023879] [Medline: 21574670]
37. Bernstein J, Heeren T, Edward E, Dorfman D, Bliss C, Winter M, et al. A brief motivational interview in a pediatric emergency department, plus 10-day telephone follow-up, increases attempts to quit drinking among youth and young adults who screen positive for problematic drinking. Acad Emerg Med 2010 Aug;17(8):890-902 [FREE Full text] [doi: 10.1111/j.1553-2712.2010.00818.x] [Medline: 20670329]
38. O'Leary TT, Monti P. Motivational enhancement and other brief interventions for adolescent substance abuse: foundations, applications and evaluations. Addiction 2004 Nov;99 Suppl 2:63-75. [doi: 10.1111/j.1360-0443.2004.00855.x] [Medline: 15488106]
39. Oetting E, Beauvais F. Peer cluster theory, socialization characteristics and adolescent drug use: A path analysis. J Couns Psychol 1987;34(2):205-213 [FREE Full text]
40. Miller W, Rollnick S. Motivational Interviewing, Second Edition: Preparing People for Change. New York: The Guilford Press; 2002.
41. Foxcroft DR, Coombes L, Wood S, Allen D, Almeida Santimano NML. Motivational interviewing for alcohol misuse in young adults. Cochrane Database Syst Rev 2014;8:CD007025. [doi: 10.1002/14651858.CD007025.pub2] [Medline: 25140980]
42. Foxcroft DR, Moreira MT, Almeida Santimano NML, Smith LA. Social norms information for alcohol misuse in university and college students. Cochrane Database Syst Rev 2015;1:CD006748. [doi: 10.1002/14651858.CD006748.pub3] [Medline: 25622306]
43. Spirito A, Sindelar-Manning H, Colby SM, Barnett NP, Lewander W, Rohsenow DJ, et al. Individual and family motivational interventions for alcohol-positive adolescents treated in an emergency department: results of a randomized clinical trial. Arch Pediatr Adolesc Med 2011 Mar;165(3):269-274 [FREE Full text] [doi: 10.1001/archpediatrics.2010.296] [Medline: 21383276]
44. Tripodi SJ, Bender K, Litschge C, Vaughn MG. Interventions for reducing adolescent alcohol abuse: a meta-analytic review. Arch Pediatr Adolesc Med 2010 Jan;164(1):85-91. [doi: 10.1001/archpediatrics.2009.235] [Medline: 20048247]
45. Tønnesen H, Ståhlbrandt H, Pedersen B. Web-based brief interventions for young adolescent alcohol and drug abusers – a systematic review. Clin Health Promot 2013:3-82.
46. Heron J, Barker ED, Joinson C, Lewis G, Hickman M, Munafò M, et al. Childhood conduct disorder trajectories, prior risk factors and cannabis use at age 16: birth cohort study. Addiction 2013 Dec;108(12):2129-2138 [FREE Full text] [doi: 10.1111/add.12268] [Medline: 23734913]
47. Vega WA, Aguilar-Gaxiola S, Andrade L, Bijl R, Borges G, Caraveo-Anduaga JJ, et al. Prevalence and age of onset for drug use in seven international sites: results from the international consortium of psychiatric epidemiology. Drug Alcohol Depend 2002 Dec 1;68(3):285-297. [Medline: 12393223]
48. Chen K, Kandel DB. The natural history of drug use from adolescence to the mid-thirties in a general population sample. Am J Public Health 1995 Jan;85(1):41-47. [Medline: 7832260]

49. Eysenbach G, CONSORT- E. CONSORT-EHEALTH: improving and standardizing evaluation reports of Web-based and mobile health interventions. J Med Internet Res 2011;13(4):e126 [FREE Full text] [doi: 10.2196/jmir.1923] [Medline: 22209829]
50. Danaher BG, Seeley JR. Methodological issues in research on web-based behavioral interventions. Ann Behav Med 2009 Aug;38(1):28-39 [FREE Full text] [doi: 10.1007/s12160-009-9129-0] [Medline: 19806416]
51. Danaher BG, Smolkowski K, Seeley JR, Severson HH. Mediators of a successful web-based smokeless tobacco cessation program. Addiction 2008 Oct;103(10):1706-1712 [FREE Full text] [doi: 10.1111/j.1360-0443.2008.02295.x] [Medline: 18715238]
52. Kelders SM, Kok RN, Ossebaard HC, Van Gemert-Pijnen JEWC. Persuasive system design does matter: a systematic review of adherence to web-based interventions. J Med Internet Res 2012;14(6):e152 [FREE Full text] [doi: 10.2196/jmir.2104] [Medline: 23151820]
53. Crutzen R, de NJ, Brouwer W, Oenema A, Brug J, de Vries NK. A conceptual framework for understanding and improving adolescents' exposure to Internet-delivered interventions. Health Promot Int 2009 Sep;24(3):277-284 [FREE Full text] [doi: 10.1093/heapro/dap018] [Medline: 19515716]
54. Arnaud N, Bröning S, Drechsel M, Thomasius R, Baldus C. Web-based screening and brief intervention for poly-drug use among teenagers: study protocol of a multicentre two-arm randomized controlled trial. BMC Public Health 2012;12:826 [FREE Full text] [doi: 10.1186/1471-2458-12-826] [Medline: 23013141]
55. Lachin JM, Matts JP, Wei LJ. Randomization in clinical trials: conclusions and recommendations. Control Clin Trials 1988 Dec;9(4):365-374. [Medline: 3203526]
56. Levy S, Sherritt L, Harris SK, Gates EC, Holder DW, Kulig JW, et al. Test-retest reliability of adolescents' self-report of substance use. Alcohol Clin Exp Res 2004 Aug;28(8):1236-1241. [Medline: 15318123]
57. Knight JR, Sherritt L, Harris SK, Gates EC, Chang G. Validity of brief alcohol screening tests among adolescents: a comparison of the AUDIT, POSIT, CAGE, and CRAFFT. Alcohol Clin Exp Res 2003 Jan;27(1):67-73. [doi: 10.1097/01.ALC.0000046598.59317.3A] [Medline: 12544008]
58. Schmiege SJ, Broaddus MR, Levin M, Bryan AD. Randomized trial of group interventions to reduce HIV/STD risk and change theoretical mediators among detained adolescents. J Consult Clin Psychol 2009 Feb;77(1):38-50. [doi: 10.1037/a0014513] [Medline: 19170452]
59. Wood MD, Fairlie AM, Fernandez AC, Borsari B, Capone C, Laforge R, et al. Brief motivational and parent interventions for college students: a randomized factorial study. J Consult Clin Psychol 2010 Jun;78(3):349-361 [FREE Full text] [doi: 10.1037/a0019166] [Medline: 20515210]
60. Saunders JB, Aasland OG, Babor TF, de la Fuente JR, Grant M. Development of the Alcohol Use Disorders Identification Test (AUDIT): WHO Collaborative Project on Early Detection of Persons with Harmful Alcohol Consumption--II. Addiction 1993 Jun;88(6):791-804. [Medline: 8329970]
61. Fishbein M, Ajzen I. Predicting and changing behavior: the reasoned action approach. New York: Psychology Press; 2010.
62. Rumpf H, Wohlert T, Freyer-Adam J, Grothues J, Bischof G. Screening questionnaires for problem drinking in adolescents: performance of AUDIT, AUDIT-C, CRAFFT and POSIT. Eur Addict Res 2013;19(3):121-127. [doi: 10.1159/000342331] [Medline: 23183686]
63. Zhang S, Paul J, Nantha-Aree M, Buckley N, Shahzad U, Cheng J, et al. Empirical comparison of four baseline covariate adjustment methods in analysis of continuous outcomes in randomized controlled trials. Clin Epidemiol 2014;6:227-235 [FREE Full text] [doi: 10.2147/CLEP.S56554] [Medline: 25053894]
64. Kypri K. Methodological issues in alcohol screening and brief intervention research. Subst Abus 2007;28(3):31-42. [doi: 10.1300/J465v28n03_04] [Medline: 18077301]
65. Berman AH, Bergman H, Palmstierna T, Schlyter F. Evaluation of the Drug Use Disorders Identification Test (DUDIT) in criminal justice and detoxification settings and in a Swedish population sample. Eur Addict Res 2005;11(1):22-31. [doi: 10.1159/000081413] [Medline: 15608468]
66. Blankers M, Koeter MWJ, Schippers GM. Missing data approaches in eHealth research: simulation study and a tutorial for nonmathematically inclined researchers. J Med Internet Res 2010;12(5):e54 [FREE Full text] [doi: 10.2196/jmir.1448] [Medline: 21169167]
67. Faul F, Erdfelder E, Lang A, Buchner A. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behav Res Methods 2007 May;39(2):175-191. [Medline: 17695343]
68. Maas CJM, Hox JJ. Robustness issues in multilevel regression analysis. Statistica Neerlandica 2004;58:127-137. [doi: 10.1046/j.0039-0402.2003.00252.x]
69. Springvloet L, Lechner L, de VH, Candel MJ, Oenema A. Short- and medium-term efficacy of a Web-based computer-tailored nutrition education intervention for adults including cognitive and environmental feedback: randomized controlled trial. J Med Internet Res 2015;17(1):e23 [FREE Full text] [doi: 10.2196/jmir.3837] [Medline: 25599828]
70. 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]
71. Little TD, Schnabel KU, Baumert J. Longitudinalmulti-group modelling with missing data. In: Modeling longitudinal and multilevel data: practical issues, applied approaches, and specific examples. Mahwah, NJ: Lawrence Erlbaum Assoc; 2000.

72. Maher C, Ferguson M, Vandelanotte C, Plotnikoff R, De BI, Thomas S, et al. A Web-Based, Social Networking Physical Activity Intervention for Insufficiently Active Adults Delivered via Facebook App: Randomized Controlled Trial. J Med Internet Res 2015;17(7):e174 [FREE Full text] [doi: 10.2196/jmir.4086] [Medline: 26169067]
73. IBM Corp. IBM SPSS Statistics for Windows, Version 21.0. In: http://www-01.ibm.com/support/docview.wss?uid=swg21476197. Armonk, NY: IBM Corp; 2012.
74. Kline RB. Principles and practice of structural equation modeling. New York: Guilford Press; 2005.
75. Pocock SJ, Assmann SE, Enos LE, Kasten LE. Subgroup analysis, covariate adjustment and baseline comparisons in clinical trial reporting: current practice and problems. Stat Med 2002 Oct 15;21(19):2917-2930. [doi: 10.1002/sim.1296] [Medline: 12325108]
76. Riper H, Spek V, Boon B, Conijn B, Kramer J, Martin-Abello K, et al. Effectiveness of E-self-help interventions for curbing adult problem drinking: a meta-analysis. J Med Internet Res 2011;13(2):e42 [FREE Full text] [doi: 10.2196/jmir.1691] [Medline: 21719411]
77. Riper H, van SA, Keuken M, Smit F, Schippers G, Cuijpers P. Curbing problem drinking with personalized-feedback interventions: a meta-analysis. Am J Prev Med 2009 Mar;36(3):247-255. [doi: 10.1016/j.amepre.2008.10.016] [Medline: 19215850]
78. Gilvarry E, McArdle P. Alcohol, drugs and young people: clinical approaches. London: Mac Keith Press; 2007:153-169.
79. Tait RJ, Hulse GK. A systematic review of the effectiveness of brief interventions with substance using adolescents by type of drug. Drug Alcohol Rev 2003 Sep;22(3):337-346. [doi: 10.1080/0959523031000154481] [Medline: 15385228]
80. Voogt CV, Kleinjan M, Poelen EAP, Lemmers LACJ, Engels RCME. The effectiveness of a web-based brief alcohol intervention in reducing heavy drinking among adolescents aged 15-20 years with a low educational background: a two-arm parallel group cluster randomized controlled trial. BMC Public Health 2013;13:694 [FREE Full text] [doi: 10.1186/1471-2458-13-694] [Medline: 23895403]
81. Barnett NP, Apodaca TR, Magill M, Colby SM, Gwaltney C, Rohsenow DJ, et al. Moderators and mediators of two brief interventions for alcohol in the emergency department. Addiction 2010 Mar;105(3):452-465 [FREE Full text] [doi: 10.1111/j.1360-0443.2009.02814.x] [Medline: 20402989]
82. Tanner-Smith EE, Steinka-Fry KT, Hennessy EA, Lipsey MW, Winters KC. Can brief alcohol interventions for youth also address concurrent illicit drug use? results from a meta-analysis. J Youth Adolesc 2015 May;44(5):1011-1023. [doi: 10.1007/s10964-015-0252-x] [Medline: 25600491]
83. Dutra L, Stathopoulou G, Basden SL, Leyro TM, Powers MB, Otto MW. A meta-analytic review of psychosocial interventions for substance use disorders. Am J Psychiatry 2008 Feb;165(2):179-187. [doi: 10.1176/appi.ajp.2007.06111851] [Medline: 18198270]
84. Palfai TP, Saitz R, Winter M, Brown TA, Kypri K, Goodness TM, et al. Web-based screening and brief intervention for student marijuana use in a university health center: pilot study to examine the implementation of eCHECKUP TO GO in different contexts. Addict Behav 2014 Sep;39(9):1346-1352 [FREE Full text] [doi: 10.1016/j.addbeh.2014.04.025] [Medline: 24845164]
85. Lee CM, Neighbors C, Kilmer JR, Larimer ME. A brief, web-based personalized feedback selective intervention for college student marijuana use: a randomized clinical trial. Psychol Addict Behav 2010 Jun;24(2):265-273 [FREE Full text] [doi: 10.1037/a0018859] [Medline: 20565152]
86. Rooke S, Copeland J, Norberg M, Hine D, McCambridge J. Effectiveness of a self-guided web-based cannabis treatment program: randomized controlled trial. J Med Internet Res 2013;15(2):e26 [FREE Full text] [doi: 10.2196/jmir.2256] [Medline: 23470329]
87. Wittchen HU, Jacobi F, Rehm J, Gustavsson A, Svensson M, Jönsson B, et al. The size and burden of mental disorders and other disorders of the brain in Europe 2010. Eur Neuropsychopharmacol 2011 Sep;21(9):655-679. [doi: 10.1016/j.euroneuro.2011.07.018] [Medline: 21896369]
88. Rohde P, Lewinsohn PM, Seeley JR, Klein DN, Andrews JA, Small JW. Psychosocial functioning of adults who experienced substance use disorders as adolescents. Psychol Addict Behav 2007 Jun;21(2):155-164 [FREE Full text] [doi: 10.1037/0893-164X.21.2.155] [Medline: 17563135]
89. Zucker RA, Donovan JE, Masten AS, Mattson ME, Moss HB. Early developmental processes and the continuity of risk for underage drinking and problem drinking. Pediatrics 2008 Apr;121 Suppl 4:S252-S272 [FREE Full text] [doi: 10.1542/peds.2007-2243B] [Medline: 18381493]
90. Apodaca TR, Longabaugh R. Mechanisms of change in motivational interviewing: a review and preliminary evaluation of the evidence. Addiction 2009 May;104(5):705-715 [FREE Full text] [doi: 10.1111/j.1360-0443.2009.02527.x] [Medline: 19413785]
91. Becker J, Haug S, Sullivan R, Schaub MP. Effectiveness of different Web-based interventions to prepare co-smokers of cigarettes and cannabis for double cessation: a three-arm randomized controlled trial. J Med Internet Res 2014;16(12):e273 [FREE Full text] [doi: 10.2196/jmir.3246] [Medline: 25486674]
92. Glasgow RE, Vogt TM, Boles SM. Evaluating the public health impact of health promotion interventions: the RE-AIM framework. Am J Public Health 1999 Sep;89(9):1322-1327. [Medline: 10474547]

93. Peters G, de BM, Crutzen R. Everything should be as simple as possible, but no simpler: towards a protocol for accumulating evidence regarding the active content of health behaviour change interventions. Health Psychol Rev 2015;9(1):1-14 [FREE Full text] [doi: 10.1080/17437199.2013.848409] [Medline: 25793484]
94. Murray E, Khadjesari Z, White IR, Kalaitzaki E, Godfrey C, McCambridge J, et al. Methodological challenges in online trials. J Med Internet Res 2009;11(2):e9 [FREE Full text] [doi: 10.2196/jmir.1052] [Medline: 19403465]
95. Babor TF, Steinberg K, Anton R, Del BF. Talk is cheap: measuring drinking outcomes in clinical trials. J Stud Alcohol 2000 Jan;61(1):55-63. [Medline: 10627097]
96. Voogt C, Kuntsche E, Kleinjan M, Poelen E, Engels R. Using ecological momentary assessment to test the effectiveness of a web-based brief alcohol intervention over time among heavy-drinking students: randomized controlled trial. J Med Internet Res 2014;16(1):e5 [FREE Full text] [doi: 10.2196/jmir.2817] [Medline: 24401555]
97. Green LW, Glasgow RE. Evaluating the relevance, generalization, and applicability of research: issues in external validation and translation methodology. Eval Health Prof 2006 Mar;29(1):126-153. [doi: 10.1177/0163278705284445] [Medline: 16510882]

Abbreviations

AUDIT-C: Alcohol Use Disorder Identification Test Consumption subscale
BAC: blood alcohol concentration
CONSORT: Consolidated Standards of Reporting Trials
EM: expectation maximization
FIML: full information maximum likelihood
ITT: intention-to-treat analyses
LMM: Linear Mixed Model
MI: motivational interviewing
RCT: randomized control trial
WDYD: What Do You Drink

Edited by G Eysenbach; submitted 10.05.15; peer-reviewed by M Blankers, S Haug; comments to author 23.07.15; revised version received 17.09.15; accepted 20.02.16; published 24.05.16.

Please cite as:
Arnaud N, Baldus C, Elgán TH, De Paepe N, Tønnesen H, Csémy L, Thomasius R
Effectiveness of a Web-Based Screening and Fully Automated Brief Motivational Intervention for Adolescent Substance Use: A Randomized Controlled Trial
J Med Internet Res 2016;18(5):e103
URL: http://www.jmir.org/2016/5/e103/
doi:10.2196/jmir.4643
PMID:27220276

Original Paper

The Electronic CardioMetabolic Program (eCMP) for Patients With Cardiometabolic Risk: A Randomized Controlled Trial

Kristen MJ Azar[1,2], MPH, MSN, RN; Suneil Koliwad[3], MD, PhD; Tak Poon[4], MD; Lan Xiao[2], PhD; Nan Lv[2], PhD; Robert Griggs[4], BA; Jun Ma[2,5,6], MD PhD

[1]Sutter Health Research, Development and Dissemination, Walnut Creek, CA, United States

[2]Palo Alto Medical Foundation Research Institute, Palo Alto, CA, United States

[3]University of California San Francisco, San Francisco, CA, United States

[4]Mills Peninsula Hospital, Burlingame, CA, United States

[5]Stanford University School of Medicine, Palo Alto, CA, United States

[6]University of Illinois, Chicago, IL, United States

Corresponding Author:
Kristen MJ Azar, MPH, MSN, RN
Sutter Health Research, Development and Dissemination
2121 N. California Blvd
Walnut Creek, CA,
United States
Phone: 1 9252874052
Email: azark@sutterhealth.org

Abstract

Background: Effective lifestyle interventions targeting high-risk adults that are both practical for use in ambulatory care settings and scalable at a population management level are needed.

Objective: Our aim was to examine the potential effectiveness, feasibility, and acceptability of delivering an evidence-based Electronic Cardio-Metabolic Program (eCMP) for improving health-related quality of life, improving health behaviors, and reducing cardiometabolic risk factors in ambulatory care high-risk adults.

Methods: We conducted a randomized, wait-list controlled trial with 74 adults aged ≥18 years recruited from a large multispecialty health care organization. Inclusion criteria were (1) BMI ≥35 kg/m^2 and prediabetes, previous gestational diabetes and/or metabolic syndrome, or (2) BMI ≥30 kg/m^2 and type 2 diabetes and/or cardiovascular disease. Participants had a mean age of 59.7 years (SD 11.2), BMI 37.1 kg/m^2 (SD 5.4) and were 59.5% female, 82.4% white. Participants were randomized to participate in eCMP immediately (n=37) or 3 months later (n=37). eCMP is a 6-month program utilizing video conferencing, online tools, and pre-recorded didactic videos to deliver evidence-based curricula. Blinded outcome assessments were conducted at 3 and 6 months postbaseline. Data were collected and analyzed between 2014 and 2015. The primary outcome was health-related quality of life. Secondary outcomes included biometric cardiometabolic risk factors (eg, body weight), self-reported diet and physical activity, mental health status, retention, session attendance, and participant satisfaction.

Results: Change in quality of life was not significant in both immediate and delayed participants. Both groups significantly lost weight and reduced waist circumference at 6 months, with some cardiometabolic factors trending accordingly. Significant reduction in self-reported anxiety and perceived stress was seen in the immediate intervention group at 6 months. Retention rate was 93% at 3 months and 86% at 6 months post-baseline. Overall eCMP attendance was high with 59.5-83.8% of immediate and delayed intervention participants attending 50% of the virtual stress management and behavioral lifestyle sessions and 37.8-62.2% attending at least 4 out of 7 in-person physical activity sessions. The intervention received high ratings for satisfaction.

Conclusions: The technology-assisted eCMP is a feasible and well-accepted intervention and may significantly decrease cardiometabolic risk among high-risk individuals.

Trial Registration: Clinicaltrials.gov NCT02246400; https://clinicaltrials.gov/ct2/show/NCT02246400 (Archived by WebCite at http://www.webcitation.org/6h6mWWokP)

(J Med Internet Res 2016;18(5):e134) doi:10.2196/jmir.5143

KEYWORDS

prevention; lifestyle; metabolic syndrome; cardiovascular disease; behavior change; health technology

Introduction

The risk for cardiometabolic diseases remains high among US adults. Although the rates of obesity, a major contributor to this risk, are leveling, up to 35% of the population continue to be classified as obese (body mass index [BMI] ≥30 kg/m^2) [1]. Addressing the prevention and treatment needs of the population remains a challenge. Effective and scalable health care systems are urgently needed to promote patient-centered population health management among patients who either have or are at risk for cardiometabolic diseases.

Lifestyle intervention integrated into ambulatory care has consistently been emphasized as a crucial approach to cardiometabolic risk reduction. In the context of obesity, lifestyle modification has been shown to produce significant risk reduction even with modest weight loss (3-5%) [2]. While intensive lifestyle interventions to reduce cardiometabolic risk have been shown to be beneficial, much work is needed to translate efficacious interventions into practical and sustainable programs that can be offered by the existing health care infrastructure. Patient-centered population health management to prevent and control cardiometabolic disease requires scalable and sustainable lifestyle interventions.

Technology-assisted approaches that are based in ambulatory care may increase the potential for widespread reach and adoption, resulting in improved long-term effectiveness and a shift towards a population based management model. Growing evidence [3-6] suggests that technology-assisted clinical tools and approaches can both increase access and decrease cost for clinic-based disease prevention and management programs that traditionally place a large burden on personnel and resources. While evidence suggests that technology-assisted lifestyle interventions for weight loss are effective, best practices remain unknown [7]. Shortcomings of existing interventions include low level of pragmatic methodology and use of technology that is not publicly available [7]. Further, there is growing concern that increased emphasis on automated online delivery modalities for lifestyle intervention can potentially fail because they sacrifice important face-to-face interactions between health care providers and patients, and among patients in group settings [8]. With this in mind, efforts to test the utility of technology-based interventions specifically for ambulatory care are underway, using widely available and low-cost tools to improve delivery mechanisms, enhance patient-provider communication, and preserve virtual face-to-face interactions while optimizing access and reach.

The purpose of this study was to examine the potential effectiveness, feasibility, and acceptability of an evidence-based group lifestyle intervention via real-time videoconferencing with other technology-assisted tools to reduce obesity and cardiometabolic risk factors among high-risk individuals in an ambulatory care setting. We hypothesized that participants would report improvements in health-related quality of life at 3 months post-baseline, compared to the delayed control group. For secondary outcomes, we hypothesized that cardiometabolic risk reduction would result in the immediate intervention group at 3 months and that these findings would be replicated in the delayed group at 6 months, while the immediate group would show continued improvement.

Methods

Study Design

The Electronic CardioMetabolic Program (eCMP) pilot study was a randomized, wait-list controlled trial among patients who either had or were at high risk for developing type 2 diabetes and/or cardiovascular disease (cardiometabolic disease). The primary end point was originally intended to be 6 months post baseline. The original intent had been for the wait-list control group to not begin the intervention until the immediate group completed the 6-month intervention, allowing for between-group comparisons at 3 and 6 months. However, this was a pilot study and logistic challenges including limited duration and funding were encountered after the study began. In order to comply with logistical limitations, we phased in the wait-list control sooner and therefore were able to compare outcomes between groups (intervention vs control) at 3 months post baseline and compare changes in outcomes within the immediate intervention group at 6 months post baseline and for the delayed-intervention group at 3 months post intervention (the mid-way point). We looked at reproducibility of early/mid-intervention effects between the first 3 months of participation and saw similar patterns of change as well as a continuation of trends.

Recruitment and Participants

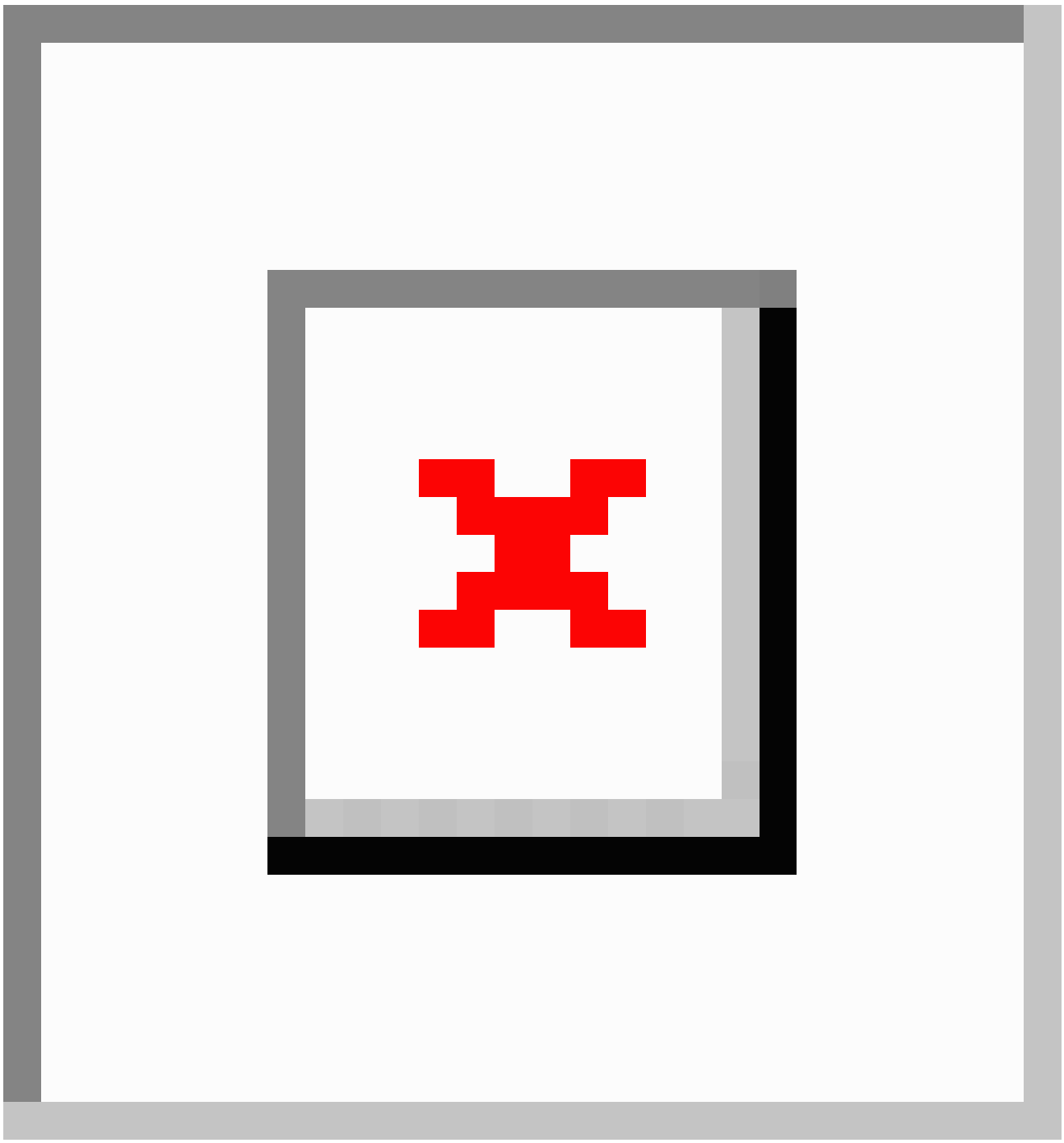
Participants were recruited from an outpatient multispecialty group practice organization in Northern California. High-risk adults (≥18 years old) in need of primary prevention of cardiovascular disease and/or diabetes (body mass index [BMI] ≥35 kg/m^2 and pre-diabetes, previous gestational diabetes and/or metabolic syndrome) or secondary prevention (BMI ≥30 kg/m^2 and type 2 diabetes and/or cardiovascular disease) were invited to participate. Participants were required to be proficient in written and spoken English and to have access to the Internet to allow for remote self-tracking, viewing of online materials and resources (eg, pre-recorded didactic videos) as well as participation in videoconference group visits. Exclusion criteria included type 1 diabetes or insulin dependence, pregnancy or active breastfeeding at the time of enrollment, current treatment for a serious medical condition (ie, cancer, except non-melanoma skin cancer), presence of any safety concerns related to significant physical or mental health issues, or life expectancy less than 12 months.

Potential eligible participants were identified through electronic health records. Participants were first screened for eligibility via phone or online and were then invited to attend the baseline assessment visit where informed consent was obtained. During the baseline visit, participants were given more information regarding the study and baseline clinical measures were obtained

and eligibility was confirmed prior to randomization. Participants were interviewed at each follow-up visit about possible adverse events during the past 3 months, and the study physician adjudicated the events per study safety protocol. Data were collected and analyzed between June 2014 and January 2015 in Burlingame, California.

The study was approved by the Palo Alto Medical Foundation Institutional Review Board. Of the 294 patients who responded to recruitment letters after their primary care provider approved study contact, 164 patients declined participation and 56 were ineligible. This process yielded the target enrollment of 74 eligible and consenting participants (see Figure 1).

Figure 1. CONSORT diagram.



Randomization and Allocation Concealment

We applied our published dynamic block randomization method [9] to assure better than chance between-treatment balance across 5 prognostic factors (age, sex, race/ethnicity, BMI, and primary prevention). The method automatically ensures allocation concealment. Participants were randomized to receive the 6-month eCMP lifestyle intervention either immediately upon randomization (n=37) or after a 3-month wait period (n=37). A designated study staff person who did not have the ability to influence the allocation system's execution performed randomization. While study group assignment was identifiable to participants and interventionists, blinding was otherwise

maintained for data collection, outcome adjudication, and data analysis.

Description of the eCMP Intervention

The goal of the eCMP is to reduce cardiometabolic risk by improving diet, physical activity, and mental health behaviors through lifestyle change. eCMP is a 6-month comprehensive program relying on weekly face-to-face group meetings via video conferencing and the delivery of evidence-based curricula using online tools and pre-recorded didactic videos presented by physicians, nutritionists, exercise physiologists, and lifestyle coaches. Technology-assisted features include (1) portable and/or wearable mobile devices for data collection (eg, Fitbit and Withings Smart Scale WS-30 [wireless scale]), (2) a comprehensive online platform and participant portal for hosting program materials (eg, homework assignments, didactic videos, and calendars), and (3) virtual small groups via real-time, encrypted, Web-based videoconferencing (see Table 1). Wireless scales were provided to the immediate intervention groups only. Technology-assisted tools for self-tracking and participation were provided for use during the study period. All participants attended an in-person orientation session post baseline assessment and prior to their first group visit. At the orientation, they participated in a technology-training workshop and received all intervention tools and materials. There were a total of 24 virtual group sessions offered alternatively between the stress management and behavioral lifestyle component and 7 in-person sessions for group-based physical activity.

Table 1. eCMP curriculum contents and delivery modalities in ambulatory care settings.

Component	Function	Features
Evidence- and theory-based curriculum[a]	Lifestyle modification and behavior change content	Weight management
		Healthy eating
		Physical activity
		Stress management
Online platform and participant portal	Hosting program materials and participant-coach communication	Didactic videos
		Homework assignments
		Calendar/schedule
		Other educational resources
Mobile monitoring devices	Participant self-monitoring, bio-feedback, and remote data capture	Wireless body scale[b]
		Pedometer
Coach-led virtual small group sessions	Curriculum content delivery for weight management, healthy eating, and stress management	Weekly (24) sessions
		Real-time, encrypted, Web-based video conferencing
Coach-led in-person sessions	Exercise curriculum content delivery	Periodic (7) sessions

[a]Curriculum used includes Diabetes Prevention Program Group Lifestyle Balance (weight management and healthy eating), Active Living Every Day (physical activity), and Heart Matters (stress management).

[b]Immediate intervention group only.

Stress Management Component

Interactive sessions for stress management consisted of biweekly virtual small group meetings of 7-10 participants. Each 1-hour session was facilitated by a trained health coach and was based on a proprietary, evidence-based [10-17] stress management curriculum called *Heart Matters*. Self-awareness is the foundation of the *Heart Matters* program, which includes these topics: belief systems, emotional responses to experiences, anger, hostility, time pressure, mindfulness, self-esteem, and forgiveness. Prior to each stress management session, participants were instructed to view an accompanying, supplemental video available via the online platform.

Behavioral Lifestyle Component

Interactive sessions for diet counseling consisted of biweekly virtual small group meetings of 7-10 participants. A registered dietitian facilitated each 1-hour session, using the core curriculum of the evidence-based Group Lifestyle Balance program developed by researchers at the University Of Pittsburgh [18,19]. Participants were asked to track their daily dietary intake via an Internet/mobile-based self-monitoring app.

Physical Activity Component

Participants were provided with instructional videos available through the online platform that encourage participants to identify opportunities throughout their day for increasing physical activity. The video content and exercise modalities were adapted from the Active Living Every Day program designed to increase physical activity among sedentary individuals [20,21]. In addition to the videos, participants were encouraged to attend 7 drop-in exercise sessions and to track their daily steps with a Fitbit pedometer.

Outcome Measures

All outcome assessors were trained to perform the measurements and interviews per standardized protocols and procedures at baseline and at 3 and 6 months post baseline. The primary outcome, overall health-related quality of life, was measured by the Short Form-8 Health Survey (SF-8), an 8-item version

of the SF-36 [22]. Secondary outcomes included biometric cardiometabolic risk factors (eg, body weight), self-reported diet and physical activity, and mental health status, retention, session attendance, and participant satisfaction. Published protocols were used to obtain height (baseline only), weight, waist circumference, and blood pressure measurements [23,24]. Participants also completed fasting blood draws at the onsite clinical laboratory for assays of plasma lipid profile. Dietary intake was assessed using multiple-pass 24-hour diet recall [25] of one typical day over the phone with participants on the Windows-based Nutrition Data System for Research (NDSR; Nutrition Coordinating Center, University of Minnesota). Diet quality was assessed using a composite Dietary Approaches to Stop Hypertension (DASH) concordance index (range 0-9) calculated by summing 9 nutrients, including total fat, saturated fat, protein, cholesterol, fiber, magnesium, calcium, sodium, and potassium [26]. Physical activity was assessed using the Stanford 7-Day Physical Activity Recall interview [27]. Stress and mental health measures included the Patient Health Questionnaire-9 (PHQ-9), Generalized Anxiety Disorder Scale (GAD-7), and Perceived Stress Scale. The PHQ-9 is a 9-item depression symptom assessment with scores ranging from 0-27. PHQ-9 scores of 5-9, 10-14, 15-19, and 20-27 represent mild, moderate, moderately severe, and severe depression, respectively [28,29]. The GAD-7 is a valid and reliable 7-question scale for screening generalized anxiety disorder and strongly associated with multiple domains of functional impairment [30]. The Perceived Stress Scale, a 14-item instrument, is a global measure of perceived stress that asks participants to report how often their lives seem to be uncontrollable or overloaded during the last year [31].

Attendance was tracked for all participants for the entire 6-month study period. Anonymous feedback and satisfaction ratings with the eCMP intervention components were obtained from participants online through survey questions ranked on a 5-point Likert scale with 1 indicating “not at all satisfied” and 5 indicating “extremely satisfied.” The survey included questions pertaining to session-related components (ie, technology, coach, and group dynamics) and home activities and materials (eg, self-monitoring activities, video/DVD, handouts, and the “virtual package”). All eCMP participants (both arms) were encouraged by the coach to complete the survey after each virtual group and/or in-person session (31 sessions total).

Statistical Analysis

Frequencies and percentages were calculated for describing categorical variables, and means and standard deviations were used for continuous variables. Intention-to-treat analyses of between-group and within-group differences in primary and secondary outcomes were tested in a repeated-measures mixed-effects linear model with adjustment of the baseline value of the outcome of interest. Between-group differences (immediate intervention vs delayed intervention) were examined at 3 months post baseline. Within-group differences for both immediate intervention and the delayed-intervention group were assessed at 6 months post baseline. All analyses were conducted using SAS, version 9.3.

Results

Retention and Baseline Characteristics

Patients were recruited from March 21, 2014, through May 13, 2014. Follow-up was completed on January 2, 2015. Of the 74 randomized participants, 89% (33/37) of intervention and 97% (36/37) of delayed intervention participants were assessed at 3 months and 86% (32/37) and 86% (32/37) at 6 months (Figure 1). There was no discernable pattern to the attrition. Participants were 59.7 (11.2) years old, mostly female, non-Hispanic white, and severely obese (see Table 2). At baseline, 27% (20/74) of participants were enrolled for primary prevention and 73% (54/74) for secondary prevention.

Table 2. Baseline characteristics.

	All (N=74)	Immediate (n=37)	Delayed (n=37)
Age in years, mean (SD)	59.7 (11.2)	59.6 (11.9)	59.8 (10.5)
Body mass index, kg/m^2, mean (SD)	37.1 (5.4)	37.0 (5.7)	37.3 (5.2)
Female, %	59.5	59.5	59.5
Race, %			
Hispanic	5.4	5.4	5.4
Non-Hispanic white	82.4	83.8	81.1
Non-Hispanic black	4.1	0	8.1
Asian/Pacific Islander	2.7	2.7	2.7
Other	5.4	8.1	2.7
Primary prevention[a], %	27.0	27.0	27.0
Secondary prevention[b], %	73.0	73.0	73.0

[a]Primary prevention is defined as BMI ≥35 kg/m^{2}and prediabetes and/or metabolic syndrome.

[b]Secondary prevention is defined as BMI ≥30 kg/m^2 and type 2 diabetes and/or cardiovascular disease.

Quality of Life, Weight, and Cardiometabolic Risk Factors

Change in health-related quality of life was not significant in both immediate and delayed participants. The mean weight loss was -2.3 kg (-2.0% of baseline) in immediate participants versus -0.2 kg (-0.01%) in delayed controls at 3 months. The between-group mean difference in change was -2.1 kg (95% CI -4.3 to 0.1) (see Table 3). At 6 months post baseline, the immediate group achieved significant weight loss of -3.1 kg (95% CI -4.7 to -1.5), achieving a 2.8% weight reduction. Equivalently, immediate intervention participants had greater reductions in BMI than delayed controls at 3 months and the net reduction was -0.7 kg/m^2 (95% CI -1.5 to -0.02). Delayed controls achieved comparable BMI reductions at 6 months. At 6 months, the immediate intervention group achieved an overall mean reduction of -1.0 kg/m^2 (95% CI -1.5 to -0.5).

Differences in waist circumference were not significant between the immediate and delayed groups at 3 months, with a decrease of -0.6 cm (95% CI -3.8 to 2.6). The immediate intervention group showed a significant reduction in waist circumference at 6 months post baseline. Participants had normal mean blood pressure, and fasting plasma lipids at baseline, with non-significant but consistent tendency of improvements in the immediate relative to the delayed group at 3 months and inconsistent changes due to insufficient sample size (n=19) for the immediate group at 6 months (see Tables 3 and 4).

Table 3. Changes in quality of life, anthropometric, blood pressure, diet, physical activity, and stress and mental health at 3 and 6 months.

		Baseline, mean (SD)	Change from baseline to 3 months, mean (95% CI)			Change from 3 to 6 months, mean (95% CI)		Change from baseline to 6 months, mean (95% CI)
		All (N=74)	Immediate (intensive phase) (n=33)	Delayed (no intervention) (n=36)	Between-group difference, difference in change	Immediate (maintenance phase) (n=32)	Delayed (intensive phase) (n=32)	Immediate (intensive+ maintenance) (n=32)
Quality of life								
	SF-8 mental component	49.2 (8.1)	-1.3 (-4.0 to 1.5)	1.5 (-1.1 to 4.1)	-2.8 (-6.6 to 1.0)	3.0 (-0.7 to 6.6)	1.3 (-2.3 to 4.8)	1.7 (-1.3 to 4.6)
	SF-8 physical component	45.1 (7.2)	2.3 (-0.4 to 5.1)	0.9 (-1.7 to 3.6)	1.4 (-2.4 to 5.2)	-2.6 (-5.8 to 0.6)	1.7 (-1.4 to 4.8)	-0.2 (-3.1 to 2.7)
Anthropometric and blood pressure								
	BMI, kg/m^2	37.1 (5.4)	-0.7 (-1.3 to -0.2)	0 (-0.5 to 0.5)	-0.7 (-1.5 to -0.02)[a]	-0.3 (-0.7 to 0.2)	-0.6 (-1.0 to -0.1)[b]	-1.0 (-1.5 to -0.5)[b]
	Weight, kg	107.4 (18.8)	-2.3 (-3.9 to -0.7)	-0.2 (-1.7 to 1.3)	-2.1 (-4.3 to 0.1)	-0.8 (-2.1 to 0.4)	-1.6 (-2.9 to -0.4)[b]	-3.1 (-4.7 to -1.5)[b]
	Weight change, %	0	-2.0 (-3.5 to -0.6)	0 (-1.4 to 1.4)	-2.0 (-4.0 to -0.1)[a]	-0.8 (-1.9 to 0.4)	-1.5 (-2.6 to -0.4)[b]	-2.8 (-4.2 to -1.4)[b]
	Waist circumference, cm	119.6 (13.9)	-1.9 (-4.2 to 0.4)	-1.3 (-3.5 to 0.9)	-0.6 (-3.8 to 2.6)	-2.2 (-4.2 to -0.2)[b]	-1.1 (-3.0 to 0.8)	-4.1 (-6.5 to -1.7)[b]
	Systolic blood pressure	124.0 (10.1)	0.7 (-2.2 to 3.6)	1.1 (-1.7 to 3.9)	-0.4 (-4.4 to 3.6)	-0.3 (-3.8 to 3.1)	-2.7 (-6.0 to 0.6)	0.3 (-2.7 to 3.4)
	Diastolic blood pressure	72.6 (10.7)	2.9 (0.6-5.3)	2.6 (0.4-4.8)	0.3 (-2.9 to 3.6)	2.0 (-1.0 to 5.0)	2.8 (-0.5 to 6.1)	5.0 (2.5 to 7.4)[b]
Diet and physical activity								
	DASH score[c]	2.4 (1.4)	0.4 (-0.1 to 0.8)	0.1 (-0.4 to 0.6)	0.3 (-0.4 to 0.9)	-0.3 (-0.9 to 0.4)	0.3 (-0.4 to 0.9)	0.1 (-0.4 to 0.6)
	Stanford 7-day Physical Activity Recall (metabolic equivalents)	680.5 (634.5)	652.8 (236.8-1068.7)	103.6 (-294.8 to 501.9)	549.2 (-26.8 to 1125.2)	-708.5 (-1224.8 to -192.2)[b]	-419.2 (-926.1 to 87.7)	-55.7 (-477.7 to 366.3)
Stress management								
	PHQ-9	5.2 (4.1)	-1.1 (-2.2 to 0)	-0.8 (-1.9 to 0.2)	-0.3 (-1.8 to 1.2)	0.3 (-0.9 to 1.6)	-0.8 (-1.9 to 0.4)	-0.8 (-1.9 to 0.4)
	GAD-7	3.4 (3.1)	0.2 (-0.7 to 1.2)	-0.5 (-1.5 to 0.4)	0.8 (-0.6 to 2.1)	-1.2 (-2.4 to 0)	0.4 (-0.8 to 1.5)	-1.0 (-2.0 to 0)
	Perceived Stress Scale	13.1 (6.1)	0.3 (-1.3 to 2.0)	-1.2 (-2.7 to 0.4)	1.5 (-0.8 to 3.8)	-1.7 (-3.9 to 0.5)	0.4 (-1.7 to 2.6)	-1.4 (-3.2 to 0.4)

[a]*P*<.05 between group difference.

[b]*P*<.05 within group difference.

[c]DASH scores were calculated based on combining nine nutrient targets (ie, total fat, saturated fat, protein, cholesterol, fiber, magnesium, calcium, sodium, and potassium). The intermediate target of each nutrient was halfway between the DASH target and population mean (based on the National Health and Nutrition Examination Surveys 2007-2008, latest data available at the inception of this study). For a nutrient, participants reaching the DASH target were assigned one point, those reaching the intermediate target were assigned a half-point, and those not meeting the intermediate target were given 0 point. The DASH score was the sum of points for all 9 nutrients.

Table 4. The changes in fasting plasma lipids by groups at 3 and 6 months.

	Baseline, mean (SD)	Change from baseline to 3 months, mean (95% CI)			Change from 3 to 6 months, mean (95% CI)		Change from baseline to 6 months, mean (95% CI)
	All (N=63)	Immediate (intensive phase) (n=22)	Delayed (no intervention) (n=32)	Between-group difference, difference in change	Immediate (maintenance phase) (n=13)	Delayed (intensive phase) (n=24)	Immediate (intensive+ maintenance) (n=19)
TC	175.1 (40.6)	-5.0 (-19.3 to 9.2)	5.4 (-5.8 to 16.7)	-10.5 (-28.7 to 7.7)	16.1 (3.1-29.1)	-0.8 (-10.4 to 8.7)	11.1 (-2.2 to 24.3)
HDL-C	49.7 (16.3)	5.0 (1.8-8.2)	2.5 (0.1- 4.9)	2.5 (-1.5 to 6.5)	2.1 (-1.3 to 5.5)	0.4 (-2.2 to 3.0)	7.2 (4.3-10.1)[a]
LDL-C	98.5 (36)	-10.7 (-23 to 1.5)	0.6 (-9.0 to 10.2)	-11.3 (-26.9 to 4.3)	16.0 (4.2-27.9)	1.8 (-7.0 to 10.6)	5.3 (-6.1 to 16.7)
TC:HDL ratio	3.8 (1.1)	-0.4 (-0.8 to 0.1)	-0.1 (-0.4 to 0.3)	-0.3 (-0.8 to 0.3)	0.2 (-0.1 to 0.4)	0 (-0.2 to 0.2)	-0.2 (-0.6 to 0.2)
Tri-glyceride	134.5 (61.5)	-0.5 (-23.8 to 22.8)	-7.1 (-28.9 to 14.6)	-12.6 (-42.5 to 17.3)	-6.7 (-28.1 to 14.8)	-14.8 (-30.5 to 0.9)	-7.1 (-28.9 to 14.6)

[a]$P<.05$ within-group difference.

Diet and Physical Activity Behaviors

DASH score and assessment of leisure-time physical activity of at least moderate intensity did not show statistically significant improvement for either group at 3 and 6 months (see Table 3).

Stress and Mental Health Measures

None of the stress and mental health measures (ie, PHQ-9, GAD-7, and perceived stress scale) had significant improvement in the immediate relative to the delayed group at 3 months (see Table 3). The mean changes in PHQ-9, GAD-7, and perceived stress scale showed non-significant but consistent tendency of improvements at 6 months for the immediate intervention group.

Intervention Attendance

Attendance at virtual and in-person group sessions varied by component. For the stress management component, attendance was higher, with 65% (24/37) attendance and 78% (29/37) attendance of immediate and delayed intervention participants, respectively, attending at least half of the 12 total sessions offered. For the behavioral lifestyle component, a majority (60%, 22/37) of immediate intervention participants attended at least half of the sessions offered and 22% (8/37) attended at least 80% of the 12 total sessions offered over the 6-month intervention. Similarly, a large majority (84%, 31/37) of delayed intervention participants attended at least half of the sessions offered, and nearly half (43%, 16/37) of them attended at least 80% of sessions offered. Further, 19% (7/37) of immediate intervention participants compared to 49% (18/37) of delayed intervention participants attended at least 80% of total sessions for stress management offered. For the physical activity component, 38% (14/37) and 62% (23/37) of immediate and delayed intervention participants, respectively, attended at least 4 of the 7 offered sessions. Adherence to 80% of sessions offered was 16% (6/37) for immediate intervention and 11% (4/37) for delayed intervention group.

Participant Satisfaction and Feedback

Among all participants (both arms) who attended sessions, 39.10% (461/1179) completed the participant satisfaction surveys. Results between the immediate and delayed intervention group were similar (see Table 5). Overall satisfaction was high with scores ranging from mean 4.1 (SD 0.9) to 4.4 (SD 0.7), with health coaches and facilitators rated as the highest satisfying component compared to other components. More than half of participants indicated that they were satisfied (42.4%, 189/446) or extremely satisfied (38.1%, 170/446) with the technology used in group sessions. Most participants were either satisfied or extremely satisfied with the coach facilitators (92.8%, 415/447) and the general group dynamics (78.7%, 350/445). Among those components, fewer than 5% of scores were rated below 3 on the Likert scale. A majority of participants indicated that they were satisfied or extremely satisfied with self-monitoring activities (78.7%, 352/447), the video or DVD resources (88.5%, 386/436), handouts (85.4%, 345/404), and the “virtual package” as a whole (77.9%, 346/444). Further, among these components, fewer than 5% of ratings fell below 3 on the Likert scale.

Table 5. Participant satisfaction results[a].

	Satisfied with the session		Satisfied with home activities and materials				
	Technology (n=446)	Coach (n=447)	Group dynamics (n=445)	Self-monitoring activities (n=447)	Video or DVD (n=436)	Handouts (n=404)	Virtual package (n=444)
Mean (SD)	4.1 (0.9)	4.4 (0.7)	4.1 (0.8)	4.1 (0.9)	4.2 (0.7)	4.2 (0.7)	4.1 (0.9)
Rating (Likert scale), %							
Not at all satisfied 1	2.0	0.2	0.7	1.1	0.7	0.2	0.5
2	4.5	0.4	4.3	4.9	1.8	2.5	6.1
3	13.0	6.5	16.4	15.2	8.9	11.9	15.5
4	42.4	41.0	45.8	38.5	49.8	51	43.7
Extremely satisfied 5	38.1	52.0	32.8	40.3	38.8	34.4	34.2

[a]Participant satisfaction was rated on a voluntary basis by participants after each virtual group or in-person session.

Adverse Events

Two hospitalizations occurred during the 6-month trial and both of them were determined by the study physician to be not related to the study. There were no deaths.

Discussion

Principal Findings

The purpose of this pilot study was to examine the potential clinical benefit, feasibility, and acceptability of a novel, evidence- and theory-based, technology-assisted behavioral lifestyle intervention for improving health-related quality of life and reducing cardiometabolic risk in ambulatory care. Our main findings suggest that although the eCMP intervention failed to improve health-related quality of life, it showed potential for decreasing cardiometabolic risk among high-risk individuals. Further, the tools and technology-assisted approaches utilized in the intervention demonstrated good feasibility and acceptability among participants.

Comprehensive lifestyle intervention has become a crucial approach to prevention and treatment of obesity [2], metabolic syndrome [32], diabetes [33,34], and cardiovascular disease [35,36]. Previous paradigms for comprehensive lifestyle intervention, which often involve at least weekly in-person one-on-one or group meetings over months to a year, are insufficient to meet the growing population health management needs of the nation. Rising prevalence and suboptimal management of cardiometabolic conditions present a major challenge to the US health care system.

These needs have risen to the national health care agenda, as the Center for Medicare and Medicaid Innovation, through legislation provided within the Affordable Care Act, has specifically called for the use of technology to improve the capacity to provide health services for patients with chronic conditions [37]. The Task Force on Community Preventive Services recommends technology-assisted, multicomponent weight-loss interventions [38]. Several recent studies have demonstrated the potential effectiveness of using digital health tools to promote behavior change and reduce cardiometabolic risk in adults [7,39-50]. Technology may provide the means by which efficacious lifestyle interventions can be translated into real-world, clinic-based settings while retaining effectiveness and increasing access and affordability. In their review, Khaylis et al identified five key components to efficacious technology-based weight loss interventions: use of a structured program, self-monitoring, feedback and communication, social support, and individual tailoring [51]. The eCMP program incorporates all of these elements. The eCMP intervention additionally incorporates several novel elements including semiremote intervention delivery, the use of virtual small groups, and the emphasis on stress management as a distinct but complementary component to diet, physical activity, and behavioral strategies.

Satisfaction with eCMP and the technology-assisted tools was overall high, suggesting good acceptability among participants. The eCMP intervention utilized virtual groups as the primary delivery modality for coach-led face-to-face interactions and curriculum delivery for the stress management and behavioral lifestyle components of the intervention. Technology-assisted interventions for weight loss that incorporate remote intervention delivery and support have been shown to produce outcomes comparable to an in-person intervention [3,40,52,53]. Attendance was generally higher in components utilizing virtual groups compared with the physical activity component that required in-person group visits. This in part may have been due to the timing of the sessions, where virtual groups were offered on weeknights and physical activity in-person sessions were offered on weekends. Technology-assisted approaches to promote behavior change, such as virtual groups, have the potential to improve adherence by making lifestyle interventions more convenient and aiding individuals in overcoming some of the barriers they may encounter in attending frequent, clinic-based, in-person sessions [53,54].

Retention rate was high in our study, compared to intervention randomized controlled trials in primary care settings [55-58]. Overall eCMP attendance was also high, relative to other studies and interventions [59,60] and especially considering the intensity

of the intervention. For the 6-month eCMP intervention, 50% attendance correlates to attendance at approximately 15 sessions in 6 months. The novel use of virtual group visits was recently shown to be effective in the delivery of weight management interventions [53,54]. Virtual group visits are a promising approach to increase accessibility of lifestyle interventions to interested individuals [54]. Participants expressed high levels of satisfaction with the technology used in group sessions, self-monitoring activities, and the "virtual package" as a whole. Participants rated group dynamics very highly, suggesting that virtual group format did not hinder the experience of being in an in-person group setting. These findings suggest that it is possible to use technology to increase the scale of an intervention without losing socially important aspects of group-based behavioral lifestyle modification that have been a crucial part of more traditional face-to-face, in-person programs. Further, the participants in this study were older, with varying levels of comfort and skill with technology use. While most participants found the training useful and were able to participate without major difficulties, it may be useful in future interventions to tailor technology training using a pre-assessment of group participants according to baseline skill level, experience, and comfort.

Measurement of health-related quality of life remained unchanged, close to the average score in the general US population [22]. Measures of clinical effectiveness and benefit were overall modest. While weight loss was modest (≤3%) among participants, reductions in waist circumference were significant in the immediate group at 6 months. Also, a majority of participants were enrolled for secondary prevention and were being medically managed for diabetes and/or cardiovascular disease (eg, hypertension and/or dyslipidemia). At baseline, blood pressure and lipid levels were well managed and near goal, making clinical effects of the intervention difficult to discern.

Limitations

Our study has a number of potential limitations. First, logistical challenges resulted in a change to the study design and primary endpoint. While this change limited our ability to analyze differences between groups after completing the intervention, we were able to assess patterns in trajectory between the two groups at similar timepoints during the course of treatment. Other limitations included a possibility of selective response to the feedback surveys among participants. Participants who responded might have been those who were more committed and more positive toward the intervention. To mitigate this possibility, all participants were highly encouraged by their coaches to complete the survey after each session and all participant evaluations were anonymous. Methodological limitations (eg, small sample size and short follow-up duration) are reflective of a pilot study. Despite these limitations, this study shows that the eCMP intervention is feasible and acceptable in a health care setting and has potential for decreasing cardiometabolic risk among high-risk patients.

Conclusion

The eCMP intervention showed potential for decreasing cardiometabolic risk among high-risk individuals and also emphasized stress management as a key component. The tools and technology-assisted approaches utilized in the intervention demonstrated good feasibility and acceptability among participants. Future interventions should continue to explore the use of technology to facilitate remote delivery of ambulatory care based interventions in order to optimize the partnership of patients and their health care providers in improving lifestyle behaviors.

Acknowledgments

This was an investigator-initiated study funded by the Palo Alto Medical Foundation. No sponsor or funding source had a role in the design or conduct of the study; collection, management, analysis, or interpretation of the data; or preparation, review, or approval of the manuscript.

The authors would like to thank Laura Nolting, BS, for her help in the administrative support of this study. We are indebted to the following individuals for their instrumental contributions to the conduct of the study: Veronica Luna, BS, Elizabeth Jameiro, MD, and Nancy Wittels, MS, and Terry Reed, RN, MS, HN-BC, Debby Hagenmaier, RN, MS, Pauline Beggs, RN, Beth Kassler, BS, Leah Groppo, RD, CDE, Dania Saarony, RD, Mishaffner Jones, RN, MS, ACSM-CES, Beryl Onedera, BS, and Tony Ferrari, BS.

Conflicts of Interest

None declared.

References

1. Ogden CL, Carroll MD, Flegal KM. Prevalence of obesity in the United States. JAMA 2014 Jul;312(2):189-190. [doi: 10.1001/jama.2014.6228] [Medline: 25005661]
2. Jensen MD, Ryan DH, Apovian CM, Ard JD, Comuzzie AG, Donato KA, American College of Cardiology/American Heart Association Task Force on Practice Guidelines, et al. 2013 AHA/ACC/TOS guideline for the management of overweight and obesity in adults: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines and The Obesity Society. J Am Coll Cardiol 2014 Jul 1;63(25 Pt B):2985-3023 [FREE Full text] [doi: 10.1016/j.jacc.2013.11.004] [Medline: 24239920]

3. Spring B, Duncan JM, Janke EA, Kozak AT, McFadden HG, DeMott A, et al. Integrating technology into standard weight loss treatment: a randomized controlled trial. JAMA Intern Med 2013 Jan 28;173(2):105-111 [FREE Full text] [doi: 10.1001/jamainternmed.2013.1221] [Medline: 23229890]
4. Cole-Lewis H, Kershaw T. Text messaging as a tool for behavior change in disease prevention and management. Epidemiol Rev 2010;32:56-69 [FREE Full text] [doi: 10.1093/epirev/mxq004] [Medline: 20354039]
5. Chaudhry B, Wang J, Wu S, Maglione M, Mojica W, Roth E, et al. Systematic review: impact of health information technology on quality, efficiency, and costs of medical care. Ann Intern Med 2006 May 16;144(10):742-752. [Medline: 16702590]
6. Free C, Phillips G, Watson L, Galli L, Felix L, Edwards P, et al. The effectiveness of mobile-health technologies to improve health care service delivery processes: a systematic review and meta-analysis. PLoS Med 2013 Jan;10(1):e1001363 [FREE Full text] [doi: 10.1371/journal.pmed.1001363] [Medline: 23458994]
7. Levine DM, Savarimuthu S, Squires A, Nicholson J, Jay M. Technology-assisted weight loss interventions in primary care: a systematic review. J Gen Intern Med 2015 Jan;30(1):107-117. [doi: 10.1007/s11606-014-2987-6] [Medline: 25134692]
8. Lyden JR, Zickmund SL, Bhargava TD, Bryce CL, Conroy MB, Fischer GS, et al. Implementing health information technology in a patient-centered manner: patient experiences with an online evidence-based lifestyle intervention. J Healthc Qual 2013;35(5):47-57. [doi: 10.1111/jhq.12026] [Medline: 24004039]
9. Xiao L, Lavori PW, Wilson SR, Ma J. Comparison of dynamic block randomization and minimization in randomized trials: a simulation study. Clin Trials 2011 Feb;8(1):59-69 [FREE Full text] [doi: 10.1177/1740774510391683] [Medline: 21335590]
10. Blumenthal JA, Sherwood A, Babyak MA, Watkins LL, Waugh R, Georgiades A, et al. Effects of exercise and stress management training on markers of cardiovascular risk in patients with ischemic heart disease: a randomized controlled trial. JAMA 2005 Apr 6;293(13):1626-1634. [doi: 10.1001/jama.293.13.1626] [Medline: 15811982]
11. Blumenthal JA, Wang JT, Babyak M, Watkins L, Kraus W, Miller P, et al. Enhancing standard cardiac rehabilitation with stress management training: background, methods, and design for the enhanced study. J Cardiopulm Rehabil Prev 2010;30(2):77-84 [FREE Full text] [doi: 10.1097/HCR.0b013e3181d0c1d3] [Medline: 20216360]
12. Cohen BE, Panguluri P, Na B, Whooley MA. Psychological risk factors and the metabolic syndrome in patients with coronary heart disease: findings from the Heart and Soul Study. Psychiatry Res 2010 Jan 30;175(1-2):133-137 [FREE Full text] [doi: 10.1016/j.psychres.2009.02.004] [Medline: 19969373]
13. Quartana PJ, Burns JW. Emotion suppression affects cardiovascular responses to initial and subsequent laboratory stressors. Br J Health Psychol 2010 Sep;15(Pt 3):511-528. [doi: 10.1348/135910709X474613] [Medline: 19840496]
14. Rozanski A, Blumenthal JA, Davidson KW, Saab PG, Kubzansky L. The epidemiology, pathophysiology, and management of psychosocial risk factors in cardiac practice: the emerging field of behavioral cardiology. J Am Coll Cardiol 2005 Mar 1;45(5):637-651 [FREE Full text] [doi: 10.1016/j.jacc.2004.12.005] [Medline: 15734605]
15. Jaremka LM, Lindgren ME, Kiecolt-Glaser JK. Synergistic relationships among stress, depression, and troubled relationships: insights from psychoneuroimmunology. Depress Anxiety 2013 Apr;30(4):288-296 [FREE Full text] [doi: 10.1002/da.22078] [Medline: 23412999]
16. Roohafza H, Sadeghi M, Naghnaeian M, Shokouh P, Ahmadi A, Sarrafzadegan N. Relationship between Metabolic Syndrome and Its Components with Psychological Distress. Int J Endocrinol 2014;2014:203463 [FREE Full text] [doi: 10.1155/2014/203463] [Medline: 24672543]
17. Bergmann N, Gyntelberg F, Faber J. The appraisal of chronic stress and the development of the metabolic syndrome: a systematic review of prospective cohort studies. Endocr Connect 2014 Jun;3(2):R55-R80 [FREE Full text] [doi: 10.1530/EC-14-0031] [Medline: 24743684]
18. Kramer MK, Kriska AM, Venditti EM, Miller RG, Brooks MM, Burke LE, et al. Translating the Diabetes Prevention Program: a comprehensive model for prevention training and program delivery. Am J Prev Med 2009 Dec;37(6):505-511. [doi: 10.1016/j.amepre.2009.07.020] [Medline: 19944916]
19. Diabetes Prevention Support Center. Group Lifestyle Balance Manual of Operations. Pittsburgh, PA: University of Pittsburgh; 2011. URL: http://www.diabetesprevention.pitt.edu/wps/wp-content/uploads/2015/01/GroupLifestyleBalanceManualOperations.pdf [accessed 2016-01-19] [WebCite Cache ID 6efEE0Xdm]
20. Wilcox S, Dowda M, Griffin SF, Rheaume C, Ory MG, Leviton L, et al. Results of the first year of active for life: translation of 2 evidence-based physical activity programs for older adults into community settings. Am J Public Health 2006 Jul;96(7):1201-1209. [doi: 10.2105/AJPH.2005.074690] [Medline: 16735619]
21. Human Kinetics. 2015. Active Living Partners URL: http://www.activeliving.info/ [accessed 2016-01-19] [WebCite Cache ID 6efEYfPXI]
22. Ware J, Kosinski M, Dewey J, Gandek B. How to Score and Interpret Single-Item Health Status Measures: A Manual for Users of the SF-8 Health Survey. Lincoln, RI: QualityMetric Incorporated; 2001.
23. PhenX (consensus measures of Phenotypes and eXposures). 2014 Feb 04. Measures from the PhenX Toolkit version February 4 2011 Ver 4.2 URL: https://www.phenxtoolkit.org/ [accessed 2016-04-28] [WebCite Cache ID 6h6q5RM3R]
24. Pickering TG, Hall JE, Appel LJ, Falkner BE, Graves J, Hill MN, Subcommittee of Professional and Public Education of the American Heart Association Council on High Blood Pressure Research. 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. Hypertension 2005 Jan;45(1):142-161 [FREE Full text] [doi: 10.1161/01.HYP.0000150859.47929.8e] [Medline: 15611362]
25. Conway JM, Ingwersen LA, Vinyard BT, Moshfegh AJ. Effectiveness of the US Department of Agriculture 5-step multiple-pass method in assessing food intake in obese and nonobese women. Am J Clin Nutr 2003 May;77(5):1171-1178 [FREE Full text] [Medline: 12716668]
26. Mellen PB, Gao SK, Vitolins MZ, Goff DC. Deteriorating dietary habits among adults with hypertension: DASH dietary accordance, NHANES 1988-1994 and 1999-2004. Arch Intern Med 2008 Feb 11;168(3):308-314. [doi: 10.1001/archinternmed.2007.119] [Medline: 18268173]
27. Blair SN, Haskell WL, Ho P, Paffenbarger RS, Vranizan KM, Farquhar JW, et al. Assessment of habitual physical activity by a seven-day recall in a community survey and controlled experiments. Am J Epidemiol 1985 Nov;122(5):794-804. [Medline: 3876763]
28. Spitzer RL, Kroenke K, Williams JB. Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. Primary Care Evaluation of Mental Disorders. Patient Health Questionnaire. JAMA 1999 Nov 10;282(18):1737-1744. [Medline: 10568646]
29. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001 Sep;16(9):606-613 [FREE Full text] [Medline: 11556941]
30. Spitzer RL, Kroenke K, Williams JBW, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med 2006 May 22;166(10):1092-1097. [doi: 10.1001/archinte.166.10.1092] [Medline: 16717171]
31. Cohen S, Kamarck T, Mermelstein R. A global measure of perceived stress. J Health Soc Behav 1983 Dec;24(4):385-396. [Medline: 6668417]
32. Grundy SM, Cleeman JI, Daniels SR, Donato KA, Eckel RH, Franklin BA, National Heart, Lung, and Blood Institute. Diagnosis and management of the metabolic syndrome: an American Heart Association/National Heart, Lung, and Blood Institute Scientific Statement. Circulation 2005 Oct 25;112(17):2735-2752 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.105.169404] [Medline: 16157765]
33. Centers for Disease Control and Prevention. Diabetes Prevention Recognition Program, Standards and Operating Procedures. Atlanta, GA: CDC; 2015. URL: http://www.cdc.gov/diabetes/prevention/pdf/dprp-standards.pdf [accessed 2016-04-28] [WebCite Cache ID 6h6qvTBWW]
34. Centers for Disease Control and Prevention (CDC). 2015. National Diabetes Prevention Program URL: http://www.cdc.gov/diabetes/prevention/lifestyle-program/index.html [accessed 2016-01-19] [WebCite Cache ID 6efEhrmH6]
35. Lloyd-Jones DM, Hong Y, Labarthe D, Mozaffarian D, Appel LJ, Van HL, American Heart Association Strategic Planning Task ForceStatistics Committee. Defining and setting national goals for cardiovascular health promotion and disease reduction: the American Heart Association's strategic Impact Goal through 2020 and beyond. Circulation 2010 Feb 2;121(4):586-613 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.109.192703] [Medline: 20089546]
36. Ebrahim S, Taylor F, Ward K, Beswick A, Burke M, Davey SG. Multiple risk factor interventions for primary prevention of coronary heart disease. Cochrane Database Syst Rev 2011(1):CD001561. [doi: 10.1002/14651858.CD001561.pub3] [Medline: 21249647]
37. American Telemedicine Association. Telemedicine in the Patient Protection and Affordable Care Act 2010. 2016. URL: http://www.americantelemed.org/docs/default-source/policy/telehealth-provisions-within-the-patient-protection-and-affordable-care-act.pdf?sfvrsn=14 [accessed 2016-01-19] [WebCite Cache ID 6efEqLprF]
38. Ostroff C, Kinicki A, Tamkins M. Organizational culture and climate. In: Borman WC, Ilgen DR, Kilmoski RJ, editors. Handbook of Psychology. New York: Wiley; 2003.
39. Mohr DC, Siddique J, Ho J, Duffecy J, Jin L, Fokuo JK. Interest in behavioral and psychological treatments delivered face-to-face, by telephone, and by internet. Ann Behav Med 2010 Aug;40(1):89-98 [FREE Full text] [doi: 10.1007/s12160-010-9203-7] [Medline: 20652466]
40. Ma J, Yank V, Xiao L, Lavori PW, Wilson SR, Rosas LG, et al. Translating the Diabetes Prevention Program lifestyle intervention for weight loss into primary care: a randomized trial. JAMA Intern Med 2013 Jan 28;173(2):113-121 [FREE Full text] [doi: 10.1001/2013.jamainternmed.987] [Medline: 23229846]
41. Sevick MA, Zickmund S, Korytkowski M, Piraino B, Sereika S, Mihalko S, et al. Design, feasibility, and acceptability of an intervention using personal digital assistant-based self-monitoring in managing type 2 diabetes. Contemp Clin Trials 2008 May;29(3):396-409 [FREE Full text] [doi: 10.1016/j.cct.2007.09.004] [Medline: 17997364]
42. Burke LE, Conroy MB, Sereika SM, Elci OU, Styn MA, Acharya SD, et al. The effect of electronic self-monitoring on weight loss and dietary intake: a randomized behavioral weight loss trial. Obesity (Silver Spring) 2011 Feb;19(2):338-344 [FREE Full text] [doi: 10.1038/oby.2010.208] [Medline: 20847736]
43. Neuhausera L, Kreps G. eHealth communication and behavior change: promise and performance. Soc Semiot 2010;20(1):9-27.
44. Krukowski RA, West DS, Harvey-Berino J. Recent advances in internet-delivered, evidence-based weight control programs for adults. J Diabetes Sci Technol 2009 Jan;3(1):184-189 [FREE Full text] [Medline: 20046664]

45. 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]
46. Neve M, Morgan PJ, Jones PR, Collins CE. Effectiveness of web-based interventions in achieving weight loss and weight loss maintenance in overweight and obese adults: a systematic review with meta-analysis. Obes Rev 2010 Apr;11(4):306-321. [doi: 10.1111/j.1467-789X.2009.00646.x] [Medline: 19754633]
47. Arem H, Irwin M. A review of web-based weight loss interventions in adults. Obes Rev 2011;12(20804523):236-243.
48. Manzoni GM, Pagnini F, Corti S, Molinari E, Castelnuovo G. Internet-based behavioral interventions for obesity: an updated systematic review. Clin Pract Epidemiol Ment Health 2011 Mar;7:19-28 [FREE Full text] [doi: 10.2174/1745017901107010019] [Medline: 21552423]
49. Spring B, Duncan JM, Janke EA, Kozak AT, McFadden HG, DeMott A, et al. Integrating technology into standard weight loss treatment: a randomized controlled trial. JAMA Intern Med 2013 Jan 28;173(2):105-111 [FREE Full text] [doi: 10.1001/jamainternmed.2013.1221] [Medline: 23229890]
50. Allen JK, Stephens J, Patel A. Technology-assisted weight management interventions: systematic review of clinical trials. Telemed J E Health 2014 Dec;20(12):1103-1120 [FREE Full text] [doi: 10.1089/tmj.2014.0030] [Medline: 25409001]
51. Khaylis A, Yiaslas T, Bergstrom J, Gore-Felton C. A review of efficacious technology-based weight-loss interventions: five key components. Telemed J E Health 2010 Nov;16(9):931-938 [FREE Full text] [doi: 10.1089/tmj.2010.0065] [Medline: 21091286]
52. Appel LJ, Clark JM, Yeh H, Wang N, Coughlin JW, Daumit G, et al. Comparative effectiveness of weight-loss interventions in clinical practice. N Engl J Med 2011 Nov 24;365(21):1959-1968 [FREE Full text] [doi: 10.1056/NEJMoa1108660] [Medline: 22085317]
53. Azar KMJ, Aurora M, Wang EJ, Muzaffar A, Pressman A, Palaniappan LP. Virtual small groups for weight management: an innovative delivery mechanism for evidence-based lifestyle interventions among obese men. Transl Behav Med 2015 Mar;5(1):37-44 [FREE Full text] [doi: 10.1007/s13142-014-0296-6] [Medline: 25729451]
54. Dunn C, Whetstone LM, Kolasa KM, Jayaratne KSU, Thomas C, Aggarwal S, et al. Using synchronous distance-education technology to deliver a weight management intervention. J Nutr Educ Behav 2014;46(6):602-609. [doi: 10.1016/j.jneb.2014.06.001] [Medline: 25052936]
55. Brown MJ, Sinclair M, Liddle D, Hill AJ, Madden E, Stockdale J. A systematic review investigating healthy lifestyle interventions incorporating goal setting strategies for preventing excess gestational weight gain. PLoS One 2012;7(7):e39503 [FREE Full text] [doi: 10.1371/journal.pone.0039503] [Medline: 22792178]
56. Cabassa LJ, Ezell JM, Lewis-Fernández R. Lifestyle interventions for adults with serious mental illness: a systematic literature review. Psychiatr Serv 2010 Aug;61(8):774-782 [FREE Full text] [doi: 10.1176/ps.2010.61.8.774] [Medline: 20675835]
57. Fleming P, Godwin M. Lifestyle interventions in primary care: systematic review of randomized controlled trials. Can Fam Physician 2008 Dec;54(12):1706-1713 [FREE Full text] [Medline: 19074715]
58. Lombard CB, Deeks AA, Teede HJ. A systematic review of interventions aimed at the prevention of weight gain in adults. Public Health Nutr 2009 Nov;12(11):2236-2246. [doi: 10.1017/S1368980009990577] [Medline: 19650959]
59. Goodpaster BH, Delany JP, Otto AD, Kuller L, Vockley J, South-Paul JE, et al. Effects of diet and physical activity interventions on weight loss and cardiometabolic risk factors in severely obese adults: a randomized trial. JAMA 2010 Oct 27;304(16):1795-1802 [FREE Full text] [doi: 10.1001/jama.2010.1505] [Medline: 20935337]
60. Vetter ML, Wadden TA, Chittams J, Diewald LK, Panigrahi E, Volger S, et al. Effect of lifestyle intervention on cardiometabolic risk factors: results of the POWER-UP trial. Int J Obes (Lond) 2013 Aug;37 Suppl 1:S19-S24 [FREE Full text] [doi: 10.1038/ijo.2013.92] [Medline: 23921777]

Abbreviations

DASH: Dietary Approaches to Stop Hypertension
eCMP: Electronic CardioMetabolic Program
GAD-7: Generalized Anxiety Disorder Scale
NDSR: Nutrition Data System for Research
PHQ-9: Patient Health Questionnaire

Edited by G Eysenbach; submitted 22.09.15; peer-reviewed by I Pires, S Mummah, S Capra, N Kaufman; comments to author 22.11.15; revised version received 22.01.16; accepted 21.02.16; published 27.05.16.

Please cite as:
Azar KMJ, Koliwad S, Poon T, Xiao L, Lv N, Griggs R, Ma J
The Electronic CardioMetabolic Program (eCMP) for Patients With Cardiometabolic Risk: A Randomized Controlled Trial
J Med Internet Res 2016;18(5):e134
URL: http://www.jmir.org/2016/5/e134/
doi:10.2196/jmir.5143
PMID:27234480

Original Paper

The Effectiveness of a Web-Based Computer-Tailored Intervention on Workplace Sitting: A Randomized Controlled Trial

Katrien De Cocker[1,2], PhD; Ilse De Bourdeaudhuij[1], PhD; Greet Cardon[1], PhD; Corneel Vandelanotte[3], PhD

[1]Department of Movement and Sports Sciences, Ghent University, Ghent, Belgium

[2]Research Foundation Flanders, Brussels, Belgium

[3]Physical Activity Research Group, School for Human Health and Social Sciences, Central Queensland University, Rockhampton, Australia

Corresponding Author:
Katrien De Cocker, PhD
Department of Movement and Sports Sciences
Ghent University
Watersportlaan 2
Ghent, 9000
Belgium
Phone: 32 264 63 23
Fax: 32 264 84 64
Email: Katrien.DeCocker@ugent.be

Abstract

Background: Effective interventions to influence workplace sitting are needed, as office-based workers demonstrate high levels of continued sitting, and sitting too much is associated with adverse health effects. Therefore, we developed a theory-driven, Web-based, interactive, computer-tailored intervention aimed at reducing and interrupting sitting at work.

Objective: The objective of our study was to investigate the effects of this intervention on objectively measured sitting time, standing time, and breaks from sitting, as well as self-reported context-specific sitting among Flemish employees in a field-based approach.

Methods: Employees (n=213) participated in a 3-group randomized controlled trial that assessed outcomes at baseline, 1-month follow-up, and 3-month follow-up through self-reports. A subsample (n=122) were willing to wear an activity monitor (activPAL) from Monday to Friday. The tailored group received an automated Web-based, computer-tailored intervention including personalized feedback and tips on how to reduce or interrupt workplace sitting. The generic group received an automated Web-based generic advice with tips. The control group was a wait-list control condition, initially receiving no intervention. Intervention effects were tested with repeated-measures multivariate analysis of variance.

Results: The tailored intervention was successful in decreasing self-reported total workday sitting (time × group: $P<.001$), sitting at work (time × group: $P<.001$), and leisure time sitting (time × group: $P=.03$), and in increasing objectively measured breaks at work (time × group: $P=.07$); this was not the case in the other conditions. The changes in self-reported total nonworkday sitting, sitting during transport, television viewing, and personal computer use, objectively measured total sitting time, and sitting and standing time at work did not differ between conditions.

Conclusions: Our results point out the significance of computer tailoring for sedentary behavior and its potential use in public health promotion, as the effects of the tailored condition were superior to the generic and control conditions.

Trial Registration: Clinicaltrials.gov NCT02672215; http://clinicaltrials.gov/ct2/show/NCT02672215 (Archived by WebCite at http://www.webcitation.org/6glPFBLWv)

(J Med Internet Res 2016;18(5):e96) doi:10.2196/jmir.5266

KEYWORDS

sedentary behavior; computer tailoring; employees; activPAL; sitting time; randomized controlled trial

Introduction

In modern societies, adults spend the majority of their waking time in sedentary behaviors, that is, activities in a sitting or reclining posture characterized by a low energy expenditure [1]. Working hours are probably the largest contributor to overall daily sitting time in adults. Flemish and Australian employees spent, respectively, about 71% and 77% of their working hours being sedentary [2-4]. Increasing evidence suggests that too much sitting is related to adverse health outcomes, independent of physical activity levels [5,6]. Both the total amount of sitting and the pattern of sitting (ie, prolonged uninterrupted periods) are associated with several adverse health effects in adults, such as obesity, metabolic syndrome, type 2 diabetes mellitus, some cancers, and all-cause and cardiovascular disease mortality [5-8]. Therefore, interventions to reduce (limit the amount of sitting) or interrupt (limit prolonged sitting bouts) sitting at work are warranted [9].

The evidence regarding the effectiveness of workplace interventions focusing on occupational sitting is growing [10,11]. Current workplace interventions are single- or multicomponent programs implementing individual or environmental strategies to influence workplace sitting. Examples of individual strategies are face-to-face behavioral change counselling sessions [12], workplace information sessions [13,14], tailored support for individual behavioral change through goal setting [13,14], and motivational interviewing [13,14]. Environmental strategies include the introduction of sit-stand workstations [15-17], portable pedal machines [18], and prompting software reminders [19,20]. Effects of these strategies are promising, with reductions in (mostly objectively measured) workplace sitting time (ranging from -28 to -262 minutes/8-hour workday) [12-14,18], and fewer (-4.6 bouts) and shorter (-5.6 minutes) sitting bouts [13].

A review of sit-stand workstations suggests that they can be effective in reducing occupational sitting time (the pooled effect size of 7 intervention studies was -77 minutes of sitting/8-hour workday), without compromising work performance [21]. However, these types of environmental interventions may not be feasible due to their high initial costs [22]. As a result, alternative interventions targeting workplace sitting are needed [9] in order to find out which other strategies are effective. In addition, most of the previous effectiveness studies were conducted in relatively small (maximum 62 participants) samples of employees or by low-quality evaluation methods [23]. This suggests that workplace intervention studies in larger samples are needed to increase our understanding of effective strategies to influence workplace sitting.

One popular public health promotion method that has been shown to successfully change a variety of health-related behaviors (dietary behaviors, alcohol consumption, smoking habits, and physical activity) is Web-based computer tailoring [24-26]. Computer-tailored interventions require participants to complete 1 or more brief assessment surveys. Then, based on participants' answers, a computer program selects the relevant feedback messages from a database (through if-then algorithms), with the intention to provide information that is as personally relevant as possible [26]. Delivered thought the Internet, this approach has several advantages, including low cost, no limitations due to time or location, and a large reach, as the Internet now has 2 billion users worldwide [27]. In addition, interactive Web-based interventions have several benefits. First, they create the opportunity for ongoing contact with, and support to, their users [28,29]. Second, they apply tools that support self-regulatory skills, such as goal-setting activities, self-monitoring tools, action planning, skill-building activities, email reminders, and booster sessions [28,29]. To our knowledge, no workplace interventions targeting sitting have examined the effectiveness of Web-based computer-tailored program, and it seems worthwhile to investigate the effects of this type of workplace interventions.

We developed a theory-driven, Web-based, computer-tailored intervention to influence sitting at work and found it to be acceptable in terms of the assessment questioning, interestingness, length, credibility, and relevance of the advice [30]. In this study, we aimed to investigate the effects of this computer-tailored intervention to influence workplace sitting [30] on objectively measured sitting time, standing time, and breaks from sitting, as well as self-reported context-specific sitting. We compared the computer-tailored advice with generic advice and a no-advice control at 1-month and 3-month follow-ups. We hypothesized that both interventions would result in beneficial changes compared with the control condition, but that the effects of the computer-tailored advice would be significantly greater than the effects of generic advice.

Methods

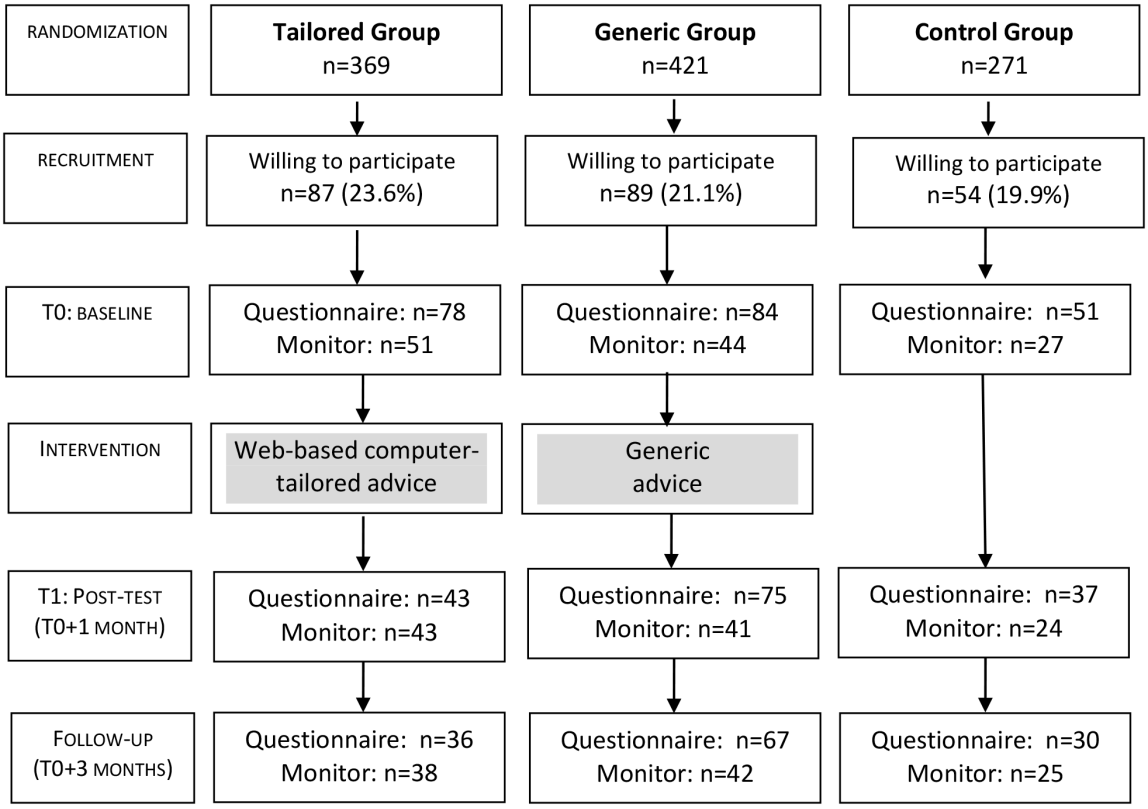
Study Design and Sample

The study used a controlled baseline (T0), 1-month follow-up (T1), and 3-month follow-up (T2: T0+3 months) design, with 3 different conditions (Figure 1 provides the study flow chart). The tailored group received a Web-based, computer-tailored intervention including personalized feedback and tips on reducing or interrupting workplace sitting (see the Interventions subsection below). The generic group received a Web-based intervention containing generic information and tips on reducing or interrupting workplace sitting. The control group was a wait-list control condition and received the generic intervention after completing all measurements.

We selected a convenience sample of 2 companies (a university and an environmental agency) in Flanders (ie, the northern, Dutch-speaking part of Belgium), mainly employing desk-based workers, having more than 100 staff members, and each having at least three different worksite locations. Both workplaces had a general health policy following European legislations and were informally committed to health aspects, but they did not yet focus on healthy sedentary behavior at work. We contacted company management by phone and email to inform them about the study, and both companies agreed to participate in this study (response rate 100%). Within the university, 3 departments of the central administration were selected to participate, and within the environmental agency, 3 departments in East Flanders were selected. Within each company, each department was randomly assigned to 1 of the 3 conditions. All selected departments were

in different physical locations and the employees had little face-to-face contact with one another, reducing the opportunity for contamination between groups.

Figure 1. Flow chart of a randomized controlled trial of an intervention to reduce and interrupt sitting time at work.



Procedures

A contact person in each department provided the email addresses of all employees. In October 2014, employees were invited to participate by email. Employees willing to participate were asked to reply to the email within 1 week, indicating whether they wanted to complete the Web-based questionnaires only or whether they wanted to complete the Web-based questionnaires and additionally wear an activity monitor. We sent a reminder email with the invitation to those who had not yet responded, 1 day before the enrollment deadline. The study protocols were approved by the Ethics Committee of the Ghent University Hospital, Belgium.

A researcher emailed a confidential website username and password to all participants. After logging in, participants received a short introduction pop-up screen and were then referred to the home page, inviting them to complete an assessment questionnaire (see the Measures subsection below). After completing the baseline questionnaire, each group received different feedback. Those who were interested in participating but who did not complete the questionnaire within 7 days were sent up to 3 automatic reminders to visit the website.

A member of the research team visited the participants willing to wear an activity monitor at their workplace on a Monday. They were instructed to wear the monitor on the thigh from Monday to Friday. A researcher covered the monitor with a transparent medical tape (Tegaderm, 3M, Diegem, Belgium) before placement, and also attached the monitor itself with this tape. These waterproof attachments allowed for 24-hour wear, including water activities (eg, bathing and swimming). Monitor data from 3 days (Tuesday, Wednesday, and Thursday) were used for this study [31]. During the 3 full days of data collection, participants were also requested to complete a day log. On Friday, the researcher recollected the monitors at the workplace. These procedures were repeated at the 3 time points (T0, T1, and T2). The study was conducted between October 2014 and March 2015.

Interventions

Computer Tailored Intervention

The development of this theory-driven intervention, called *Start to Stand*, has been described in detail elsewhere [30]. In brief, in this Web-based intervention, after completing an assessment questionnaire, users received personalized computer-tailored feedback about their sitting time, including tips and suggestions on how to interrupt (taking short standing breaks) and reduce (replacing sitting by periods of standing) sitting. Several questions on the Web-based questionnaire (see the Measures subsection below) were used to build up the feedback. The information requested from the participants referred to job-related information, knowledge about sedentary behavior, level of sitting time in 5 domains, frequency of interruptions in prolonged sitting, and level of physical activity. A set of decision rules selected the feedback messages that were matched and tailored to the specific answers given by the users. The combination of these feedback messages formed the tailored advice that appeared immediately on the user's screen. We took

into account constructs of the self-determination theory [32], the theory of planned behavior [33], and self-regulation theory [34]. Web-based interventions that use theory are more likely to have positive effects than those that are not theory based [35]. Interventions tailored to 4 or 5 theoretical concepts, such as social support, self-efficacy, and stages of change, were found to be more positive than interventions tailored to fewer than 3 theoretical concepts [36].

To increase their knowledge, participants first received general information about the importance of sitting behavior to improve health outcomes. This was followed by normative feedback about their own sitting behavior on working and nonworking days, in order to increase awareness of participants' levels of sitting time. Further, feedback on the frequency of breaks from sitting, information on the importance of these breaks for health, and the suggestion to interrupt prolonged sitting every 30 minutes was provided [14,37]. Also the negative health impact of too much sitting, independent of physical activity levels, was explained. Next, feedback on participants' physical activity level was given (see [30] for details).

At the end of this advice (section 1), participants were able to request up to 5 other noncommittal sections if they were interested. In line with the self-determination theory [32], users were able to choose which section of the advice they wanted to request additionally, instead of forcing users to complete all assessment questions at once and providing extensive advice containing all information. Leaving a choice for the users is assumed to increase autonomous motivation, which is more likely to lead to behavioral change [32]. The focus of these additional sections was on sitting during work hours (section 2: standing breaks; section 3: replacing sitting by standing), sitting during commuting to work (section 4), sitting during lunch breaks at work (section 5); and on making an action plan to improve sitting behavior (section 6). For each section, a brief questionnaire assessed participants' current sitting behaviors and related psychosocial correlates (attitudes, self-efficacy, social support, intentions, and perceived benefits and barriers) [30]. Throughout the advice, the tailoring constructs were based on the theory of planned behavior [32], a theory focusing on the intention (or motivation) to adopt or modify a behavior and assuming that the intention is higher when people have more positive attitudes, perceived social influence, and self-efficacy. After participants completed each assessment questionnaire, the tailored advice appeared immediately on the computer screen. Participants were given personalized feedback about their attitudes, self-efficacy, social support, knowledge, intentions, and perceived benefits and barriers related to reducing or interrupting sitting.

In the last section, participants who were motivated to change their sitting were invited to create an action plan to convert intentions into specific actions through the specific, measurable, attainable, relevant, and time-bound (SMART) goals and implementation intentions [34,38], as goal setting has been found to be a common and effective intervention technique used in other health behavioral change programs. The action plan operates within the self-regulation theory, targeting pre- and postintentional processes to guide behavioral change [34]. Users were asked *what* (increase standing breaks, or replace sitting by standing, or both) they wanted to do, *how long*, *how often*, *when*, and *how* in order to state personally relevant and attainable goals. Finally, users were able to select precomposed if-then statements to make an if-then plan. Users were also able to formulate if-then statements themselves in an open-ended question format. When these questions were completed, a schematic overview of this personalized action plan was immediately provided on the screen [30].

Some screenshots of the tool are provided in Multimedia Appendix 1.

Generic Intervention

In the generic advice condition, after completing the baseline assessment questionnaire, users received generic information on the importance of reducing and interrupting sitting, and generic tips and suggestions on how to interrupt (taking short standing breaks, 6 tips) and reduce (replacing sitting by periods of standing, 8 tips) sitting during work hours, (lunch) breaks, and commuting (topics similar to the tailored group). While the information covered the same topics as in the tailored group, the generic group didn't receive personalized advice or an action plan, and all information appeared on a single screen page.

Measures

Web-Based Questionnaire

The questionnaire consisted of several parts, and all questions were asked at T0, T1, and T2, except for unchangeable variables, such as height.

Sociodemographics

The following sociodemographic variables were assessed: sex, age, highest educational degree with 5 options dichotomized into low (no diploma, elementary school, secondary school) and high (high school, university) education, height, and weight.

Work-Related Variables

We asked about the number of workdays per week (1–7), average daily amount of time (hours and minutes) spent at the workplace (open-ended), occupational status (blue collar, white collar, management), and employment duration (14 categories ranging from 1–6 months to 55–60 years).

Self-reported Behavioral Measures

We assessed the level of sitting time in 5 domains using the Workforce Sitting Questionnaire (WSQ) [39]. This questionnaire assessed time spent sitting on a workday and a nonworkday for the last 7 days while (1) travelling to and from places, (2) being at work, (3) watching television, (4) using a computer at home (not work related), and (5) doing other leisure activities. The WSQ has acceptable reliability (interclass correlation coefficient=.63) and validity against objectively accelerometer-measured sitting time (r=.34 to r=.45) [39]. Participants were also asked to report the average number of breaks from sitting they have on a regular workday [40].

The validated International Physical Activity Questionnaire (IPAQ) short version [41,42] assessed the number of days and duration of time spent in walking, moderate-intensity physical

activity, and vigorous-intensity physical activity in the preceding week.

Activity Monitor and Day Log

Sedentary behavior was measured objectively using the activPAL (PAL Technologies, Glasgow, UK) activity monitor (weight 15 g, dimensions 53 × 35 × 7 mm). This inclinometer, distinguishing periods of sitting or lying from standing and assessing breaks from sitting, has been validated (correlation between activPAL and direct observation: R^2=.94) in adults [43] and was recommended to assess sedentary behavior [43,44].

Participants were requested to complete a day log and record the type of day (workday at home, workday at the workplace, or nonworkday), time of getting up, the start and end time of the working day, and the time of going to sleep.

Website Usage Statistics

We collected the number of participants requesting the different sections of the advice from the website administration. Google Analytics provided data on website visiting time [45].

Data Reduction

Data recorded by activPAL were reduced using PAL Technologies software (version 6.4.1). We calculated waking sitting time by subtracting sleep time reported in the day log from the total sitting time recorded by the activPAL device. The percentage of working time spent sitting was calculated as sitting time during work hours/work hours × 100. We used similar formulas to calculate percentage of working time spent standing and number of breaks per working hour. We calculated average values for T0, T1, and T2 from the mean scores of the 3 measurement days.

Within the 5 domains assessed using the WSQ, we truncated values over 12 hours/day to 12 hours to avoid unrealistic values [30]. We calculated total time spent sitting on a workday and on a nonworkday by summing the time reported in every domain. These totals were truncated at 16 hours/day [30].

Based on the guidelines for data processing and analysis of the IPAQ [46], total scores for walking, and moderate and vigorous physical activities were computed (number of days × time duration). Body mass index (BMI) was calculated as weight in kg divided by height in m^2.

Data Analyses

All analyses were conducted in SPSS version 22.0 (IBM Corporation) and significance was set at P<.05, while P values between .05 and .10 were interpreted as indicating borderline significance. We did not perform any multilevel analyses because there were fewer than 10 clusters and none of the variables were measured at cluster level [47]. The 3 groups' characteristics at baseline were compared using 1-way analysis of variance (for the quantitative variables age and BMI) or χ^2 tests (for the qualitative variables sex, education, work regimen, occupational status, and employment duration). We conducted the same analyses to compare the characteristics of participants willing to wear the activity monitor at baseline and those not willing to wear it. Descriptive statistics for usage of the various website components are provided.

To investigate the 1-month and 3-month follow-up effects of the intervention, we conducted 3 repeated-measures multivariate analysis of variance tests with time (T0, T1, or T2) as the within-participants factor, condition (3 groups) as the between-participants factor, and self-reported sitting (workday and nonworkday total sitting; average daily domain-specific sitting) and objectively measured sitting (total waking sitting time, working time spent sitting, working time spent standing, breaks from sitting per work hour) as the dependent variables. When the time (3 levels) × condition (3 levels) effects were significant, we conducted additional post hoc repeated-measures analyses, including 2 times points (T0–T1 or T0–T2) and only 2 conditions, to find out where the differences in changes over time between the conditions occurred. We included the following covariates in the analyses: age, sex, education, hours at work, employment duration, BMI, walking, and moderate and vigorous-intensity physical activity at baseline. Due to the skewed nature of the outcomes, we did the analyses on square root transformations to improve normality, but for reasons of clarity, we report nontransformed average scores in the tables. We executed this approach using both a retained sample analysis (ie, completer analysis) and an intent-to-treat analysis (last value carried forward). Because we found no differences between the 2 analyses, we report results only of the retained sample analysis.

Results

Participants' Characteristics at Baseline and Website Usage

The total sample (N=213) of employees completing the Web-based questionnaire consisted of 31.5% (67/213) men, 81.7% (174/213) with a high level of education, 91.5% (195/213) who were white collar workers, and 69.5% (148/213) with an employment duration of more than 5 years. Participants had a mean age of 40.3 (SD 9.1) years, worked on average 8.0 (SD 0.7) hours/day, and had a mean BMI of 23.9 (SD 3.4) kg/m^2. Table 1 gives an overview of the baseline variables for the 3 intervention groups. The 3 groups did not differ in sociodemographic, work-related, and health-related variables (see Table 1). Self-reported and objectively sitting variables at baseline are presented below. Of those completing the questionnaires, 122/213 (57.3%) employees were willing to wear the activity monitor at baseline. There were no differences between participants who were willing and those who were not willing to wear the activity monitor across all variables (sociodemographic, work-related, health-related, and self-reported sitting time variables; data not shown).

Table 1. Baseline characteristics for the 3 study groups.

Variables	Tailored group (n=78)	Generic group (n=84)	Control group (n=51)	Group comparisons	
				F_{df} or χ^2_{df}	*P* value
Sociodemographic variables					
Age in years, mean (SD)	40.5 (8.6)	40.7 (9.7)	39.3 (9.0)	$F_{2,209}$=0.44	.65
Males, n (%)	25 (32.1)	27 (32.1)	15 (29.4)	$\chi^2_{2,212}$=1.9	.76
High school/university education: n (%)	58 (75.3)	70 (83.3)	46 (90.2)	$\chi^2_{2,212}$=4.8	.09
Work-related variables					
Hours at work, mean (SD)	8.0 (0.9)	8.0 (0.6)	8.0 (0.6)	$F_{2,211}$=0.36	.70
White collar occupational status, n (%)	74 (96.1)	75 (89.3)	46 (90.2)	$\chi^2_{2,212}$=2.8	.24
Employment duration>5 years, n (%)	55 (71.4)	56 (66.7)	37 (72.5)	$\chi^2_{2,212}$=0.7	.72
Health-related variables					
BMI[a] in kg/m^2, mean (SD)	24.2 (3.1)	23.6 (3.5)	23.7 (3.5)	$F_{2,211}$=0.75	.48
Walking time in minutes/day, mean (SD)	18.8 (28.3)	21.1 (21.6)	18.6 (19.0)	$F_{2,208}$=0.30	.74
Moderate-intensity PA[b] in minutes/day, mean (SD)	24.7 (26.9)	19.3 (20.1)	18.0 (19.0)	$F_{2,208}$=1.69	.19
Vigorous-intensity PA in minutes/day, mean (SD)	8.4 (11.5)	11.6 (15.6)	9.9 (15.6)	$F_{2,209}$=1.05	.35

[a]BMI: body mass index.

[b]PA: physical activity.

At baseline, all 78 participants in the tailored group completed section 1 (100%). The average time needed to complete the assessment survey was 16.3 minutes. Time spent on the first advice page was on average 20.1 minutes. A total of 66/78 participants completed section 2 (84.6%), 64/78 completed section 3 (82.1%), 60/78 completed section 4 (76.9%), 59/78 completed section 5 (75.6%), and 54/78 completed an action plan (69.2%).

Intervention Effects on Self-reported Sitting Measures

Table 2 presents the baseline, 1-month, and 3-month follow-up values of the self-reported sitting for each group. Total workday sitting change from baseline to follow-up was significantly different between the 3 groups (see Table 2). The decrease in sitting time in the tailored group was significantly greater than the decrease in the generic group (*P*=.002) and the increase in the control group (*P*=.002). The decrease in the generic group was borderline significantly different from the change in the control group (*P*=.05). The changes over time in total nonworkday sitting did not differ significantly between the groups (see Table 2).

Table 2. Mean self-reported sitting at baseline (T0), 1-month follow-up (T1), and 3-month follow-up (T2) for the 3 groups and time × group effects.

Group	T0	T1	T2	$F_{dftime \times group}$	*P* value
Total sitting in minutes/day, mean (SD)					
Total workday sitting				T0–T1–T2: $F_{4,128}$=5.65	<.001***
Tailored (n=36)	507 (104)	480 (128)	425 (110)	T0–T1: $F_{2,149}$=1.45	.24
Generic (n=64)	457 (107)	444 (105)	437 (95)	T0–T2: $F_{2,128}$=8.47	<.001***
Control (n=28)	449 (126)	434 (131)	469 (92)		
Total nonworkday sitting				T0–T1–T2: $F_{4,128}$=1.20	.31
Tailored (n=36)	141 (70)	139 (69)	132 (70)	T0–T1: $F_{2,149}$=0.63	.54
Generic (n=64)	130 (63)	131 (67)	141 (77)	T0–T2: $F_{2,128}$=1.15	.32
Control (n=28)	123 (58)	117 (47)	134 (55)		
Domain-specific sitting in minutes/day, mean (SD)					
Sitting at work[a]				T0–T1–T2: $F_{4,118}$=6.72	<.001***
Tailored (n=33)	338 (107)	279 (92)	259 (88)	T0–T1: $F_{2,138}$=12.5	<.001***
Generic (n=61)	288 (59)	279 (64)	280 (69)	T0–T2: $F_{2,119}$=10.09	<.001***
Control (n=24)	281 (65)	280 (50)	288 (48)		
Sitting during transport[b]				T0–T1–T2: $F_{4,118}$=0.45	.77
Tailored (n=33)	78 (84)	103 (124)	58 (49)	T0–T1: $F_{2,138}$=0.47	.63
Generic (n=61)	66 (79)	60 (67)	48 (31)	T0–T2: $F_{2,119}$=0.01	.98
Control (n=24)	81 (106)	74 (88)	62 (62)		
Television viewing[b]				T0–T1–T2: $F_{4,118}$=1.31	.23
Tailored (n=33)	100 (57)	104 (56)	106 (61)	T0–T1: $F_{2,138}$=1.24 0.29	.29
Generic (n=61)	95 (62)	92 (67)	102 (68)	T0–T2: $F_{2,119}$=1.24	.29
Control (n=24)	91 (68)	79 (68)	82 (61)		
Personal computer use[b]				T0–T1–T2: $F_{4,118}$=1.51	.20
Tailored (n=33)	50 (46)	53 (47)	47 (29)	T0–T1: $F_{2,138}$=0.20	.82
Generic (n=61)	51 (40)	52 (44)	51 (39)	T0–T2: $F_{2,119}$=1.35	.26
Control (n=24)	58 (62)	59 (71)	69 (65)		
Other leisure time sitting[b]				T0–T1–T2: $F_{4,118}$=1.86	.12
Tailored (n=33)	101 (42)	90 (44)	75 (32)	T0–T1: $F_{2,138}$=1.68	.19
Generic (n=61)	99 (61)	98 (61)	97 (46)	T0–T2: $F_{2,119}$=3.64	.03*
Control (n=24)	95 (48)	96 (43)	102 (64)		

[a]Average on workday.

[b]Average of workday and nonworkday.

* *P*<.05; *** *P*<.001.

Analyses of the domain-specific sitting data showed that changes over time in sitting at work and other leisure time sitting differed significantly between the 3 groups. There was a decrease in sitting at work in the tailored group, which was significantly greater than the changes in the generic group (T0–T1: *P*<.001, T0–T2: *P*<.001) and the control group (T0–T1: *P*=.001, T0–T2: *P*=.001). The changes over time did not differ significantly between the generic group and the control group (T0–T1: *P*=.48, T0–T2: *P*=.26). There was also a decrease in other leisure time sitting from baseline to follow-up in the tailored group, which was significantly greater than the changes over time in the generic group (*P*=.007) and the control group (*P*=.02). The changes from baseline to follow-up did not differ significantly between the generic group and the control group (*P*=.78). The changes over time in sitting during transport, television viewing, and computer use did not differ significantly between the 3 groups (see Table 2).

Intervention Effects on the Objectively Measured Variables

Table 3 gives an overview of the sedentary behavior measures derived from the activPAL at baseline, and 1-month and 3-month follow-ups for each group. The changes over time in total sitting while being awake, sitting at work, and standing at work did not differ significantly between the 3 groups (see Table 3). For breaks at work, the only borderline significant difference was found between the tailored group (slight increase in breaks from baseline to follow-up) and the generic group (slight decrease in sitting time from baseline to follow-up, P=.07).

Table 3. Mean objectively measured variables at baseline (T0), 1-month follow-up (T1), and 3-month follow-up (T2) for the 3 groups and time × group effects.

Group	T0	T1	T2	$F_{\text{dftime x group}}$	P value
Total sitting time awake in hours/day, mean (SD)				T0–T1–T2: $F_{4,75}$=0.56	.69
Tailored (n=35)	576 (109)	600 (91)	607 (117)	T0–T1: $F_{2,80}$=0.52	.60
Generic (n=35)	578 (101)	574 (103)	576 (109)	T0–T2: $F_{2,79}$=1.06	.35
Control (n=23)	605 (96)	616 (115)	623 (100)		
Sitting at work in % work hours, mean (SD)				T0–T1–T2: $F_{4,75}$= $F_{4,75}$=0.22	.93
Tailored (n=35)	66.8 (15.5)	71.7 (14.0)	69.0 (13.7)	T0–T1: $F_{2,80}$=0.28	.76
Generic (n=35)	69.0 (13.8)	71.2 (15.1)	68.8 (15.1)	T0–T2: $F_{2,79}$=0.12	.89
Control (n=23)	74.3 (15.5)	78.3 (11.1)	74.8 (13.5)		
Standing at work in % work hours, mean (SD)				T0–T1–T2: $F_{4,75}$=0.10	.98
Tailored (n=35)	24.7 (13.5)	22.2 (9.0)	23.6 (11.7)	T0–T1: $F_{2,80}$=0.05	.95
Generic (n=35)	24.4 (11.3)	22.7 (15.4)	24.3 (14.4)	T0–T2: $F_{2,79}$=0.11	.90
Control (n=23)	16.3 (9.3)	17.1 (7.9)	17.8 (9.0)		
Breaks at work in no/work hour, mean (SD)				T0–T1–T2: $F_{4,75}$=2.54	.09*
Tailored (n=35)	3.8 (1.5)	3.7 (1.3)	4.3 (1.6)	T0–T1: $F_{2,80}$=0.72	.40
Generic (n=35)	3.6 (1.3)	3.6 (1.4)	3.5 (1.3)	T0–T2: $F_{2,79}$=2.40	.11
Control (n=23)	3.0 (1.4)	3.2 (1.4)	3.3 (1.6)		

* P<.10

Discussion

To our knowledge, this is the first randomized controlled study evaluating 1-month and 3-month follow-up effects of a theory-driven, Web-based, computer-tailored intervention to reduce or interrupt sitting among employees. Results are promising, with positive intervention effects on self-reported sitting time at work, self-reported sitting time during leisure, and objectively measured breaks at work. For these outcomes, the tailored intervention had superior effects to those of the control and the generic condition, confirming our hypothesis. This suggests the significance of computer tailoring in targeting sedentary behavior, as also seen in Web-based advice for other health-related behaviors, such as physical activity and diet [24-26]. However, the expectation that the effect of the generic condition would differ from the control condition was not confirmed. The provision of nontailored Web-based advice seems not sufficient to result in sedentary behavioral change, which is in contrast to behavioral change interventions targeting physical activity or dietary habits, in which generic interventions did have positive effects [48,49].

It should be noted that the positive findings concerning the decrease in self-reported sitting duration were not reflected in the objective measures, as no effect was found on activPAL-measured total sitting time or sitting time at work. This result emphasizes the importance of combining self-reported and objective measures. The effectiveness study of sit-stand workstations conducted by Chau et al [50] found the opposite pattern, namely positive effects for sitting time at work measured by the activPAL, but no effects for sitting time at work measured with the WSQ. The review of Prince et al [9] found no significant effect differences between self-reported and objectively measured sedentary time. It should be noted that most studies of the review used one of two types of measures: either objective or self-reported measures, but not both. When looking for potential explanations for the discrepancy between our self-reported and objectively measured data, we conducted post hoc correlation analyses (r=.07 for total sitting and r=.11 for work-related sitting), which suggested that the perception of self-reported sitting was different from objectively measured sitting. Some other studies showed that participants underestimated their self-reported sedentary behavior [51]. In contrast, other validation studies of the WSQ, using accelerometers, showed correlates of .34 and more [39]. Based on our study and earlier studies [51], the combination of self-reported and objective measurements is recommended to fully explore findings. A second possible explanation may be

that the sample of participants providing the activPAL data were only a subsample (122/213, 57.3% of the total sample at baseline, 105/133, 78.9% at 3-month follow-up). Still, intervention effects on the self-reported measures were not different between employees wearing the monitor and those who did not (data not shown). Even though the subsample is quite large for a field-based study, objective data should ideally have been available for all participants, as the observed statistical power was low to find differences between groups in total sitting and in sitting at work (observed power ranged from .12 to .36) and in standing time at work (observed power=.06).

Compared with baseline, the self-reported work-related sitting time in the tailored condition was lower at 1-month follow-up (-59 minutes/day) and 3-month follow-up (-79 minutes/day), while this was not the case for the other conditions. These reductions in sitting time are similar to those seen in interventions implementing activity-permissive workstations (-77 minutes/8-hour workday) as shown in the review of Neuhaus et al [21], which included 38 studies. Our reductions in sitting were even higher than those found in a meta-analysis of 34 lifestyle intervention studies (including 8 workplace interventions) having sedentary behavior as an outcome, showing an overall reduction in sedentary time by mean differences of -22 minutes/day in favor of the intervention group [52]. A reduction in sedentary time of just 30 minutes/day is suggested to have clinically meaningful effects on health [9], which indicates that our computer-tailored intervention may potentially be useful in public health promotion. However, in the review of Martin et al [53], the authors concluded that it is not known whether the effective interventions aimed at reducing sedentary behavior resulted in clinically meaningful and sustained improvements in health outcomes, as it was not possible to determine the intervention effect of reduced sedentary behavior on cardiometabolic risk, body composition, and mental health outcomes. Still, in a Spanish Internet-delivered workplace intervention focusing on decreasing occupational sitting (via goal setting for step counts and walking), occupational sitting time (-22 minutes/day) and waist circumference (-0.8 cm) decreased in the intervention group [54].

In our subsample wearing the activPAL, we found a positive intervention effect at follow-up for breaks during work (+0.4 breaks/work hour, ~3.2 breaks in an 8-hour workday, an increase from 30 to 34 breaks in an 8-hour workday) in the tailored condition compared with no change in the generic condition. In the review of Martin et al [53], pooled intervention effects on sedentary behavior patterns indicated no statistically significant effect for the number of sitting breaks per hour. The clinical meaning of the effect we observed is, however, unknown, as the dose-response relation regarding breaks is unclear [52]. Still, an eHealth intervention to reduce prolonged occupational sitting that passively prompted desk-based employees every 45 minutes to stand and perform nonexercise physical activity did examine the clinical effects. The study resulted in activity breaks about 6 times/workday for about 1.3 minutes at a time in the experimental group and showed that the mean arterial pressure significantly decreased [55] and the work-related energy expenditure increased [56] in the experimental group compared with a control group. Given that we found an increase in breaks in our study, we can assume that this computer-tailored intervention can also result in significant health effects; however, we did not track the duration of the activity breaks.

It may be surprising that an individual-based intervention such as our computer-tailored intervention resulted in such a relatively high reduction in self-reported sitting time at work. In the case of implementing sit-stand workstations, it is reasonable that sitting time is substantially reduced, as the environment is changed to do so, but without standing desks one could expect the intervention to have less effect on total sitting duration and more on the sitting pattern. Based on the feasibility and acceptability study of our computer-tailored intervention, pointing out that employees perceived interrupting sitting to be more achievable than reducing workplace sitting, this could also be expected to be the case here [30]. Nevertheless, our intervention seemed to affect both the pattern (objectively measured data) and the duration (self-reported data) of sitting.

We found no other 1-month follow-up effects on self-reported outcomes (sitting during transport, television viewing, computer use, leisure time sitting) or other objectively measured outcomes (total sitting time, sitting and standing time at work). However, we did find significant and positive 3-month follow-up intervention effects for self-reported total workday sitting and self-reported leisure time sitting. The change over time was more positive in the tailored condition than in the other conditions. The fact that leisure time sitting decreased (-26 minutes/day) from baseline to follow-up was surprising, as the advice mostly focused on work-related aspects: work hours, commuting, and (lunch) breaks. Still, this may mean that employees transferred the information and tips regarding one specific setting (work) to another (leisure). The study of Chau et al [50] obtained a similar result, showing that self-reported television viewing time decreased after a workplace intervention in which employees used a sit-stand workstation for 4 weeks. The authors argued that a reduction in television viewing time would be a welcome side effect of their intervention, especially because compensatory effects for occupational sitting (ie, less sitting at work would lead to more sitting at home) were not found in other previous studies [57-59]. In our study, however, we found no effect on television viewing.

Study Strengths and Limitations

Our study has several strengths and limitations to take into account. The first strength is the randomized controlled design with a large sample of employees relative to other workplace interventions focusing on sedentary behavior. Second, the use of the activPAL as an objective measure for the outcomes was a strength, as self-reported measures can have recall and social desirability biases. However, as stated earlier, the first limitation is that we used this monitor only in a subsample of employees willing to wear the monitor, which is probably a result of the field-based approach of this study. Further, from a methodological point of view, it would have been more suitable to randomly allocate the monitors within the total sample in order to avoid sampling bias. However, from a compliance point

of view, we believed it was better to provide a monitor only to those willing to wear one, in order to limit dropout. The dropout rate was lower (17/122, 13.9%) in the group wearing a monitor (and completing the questionnaires) than in the sample only completing the questionnaires (80/213, 37.6%). A study on the retention rates in physical activity interventions in workplace, health care, and home- or community-based settings revealed a mean retention rate of 78%, with minimal differences between intervention settings [60]. In our study, the mean retention rate was 74.3% (only questionnaire: 133/213, 62.4%; questionnaire and monitor: 105/122, 86.1%). The second limitation, probably resulting from the field-based approach, is the relatively low initial response rate (230/1061, 21.5%), which is comparable with another workplace intervention study using only 1 mailing to recruit (response rate 20%) [61]. The study of Waters et al [60] showed that response rates in physical activity interventions in adults range from 20% to 89%. Third, we recruited participants in only 2 companies, with worksites in 3 different settings, probably resulting in different workplace cultures. In addition, the study sample consisted of mainly healthy weight, highly educated women. This is in line with other workplace physical activity interventions in which participants were found to be younger, more educated, and healthier than nonparticipants [60]. All these factors may compromise generalizability of the results. Fourth, we do not know the long-term effects of this intervention, as we completed the follow-up measures only 3 months after baseline. Martin et al [53] found that interventions of up to 3 months resulted in a significant reduction in sedentary time (-48 minutes/day) in favor of the intervention group, whereas longer intervention durations of more than 3 months did not show beneficial intervention effects.

This study opens perspectives for future research. The effect of this intervention on psychosocial correlates should be tested, including the mediating effect of the change in these factors on the behavioral effects. As it stands, we do not know what the active intervention components are. Further, future research should investigate whether this tailored intervention would be more effective in combination with other (environmental) strategies, for example the use of sit-stand desks. Previous interventions also chose multicomponent programs to tackle the problems of too much sitting [50]. Therefore, this intervention could be an additional component in studies based on ecological models, intervening simultaneously at multiple levels.

Conclusions

To our knowledge, this is the first intervention study to describe the effectiveness of a theory-driven, Web-based, interactive computer-tailored intervention aimed at reducing and interrupting sitting at work. The computer-tailored approach showed promising outcomes to address sitting time, as the tailored intervention was successful in decreasing self-reported sitting time at work and during leisure time, and in increasing objectively measured breaks at work compared with the generic and control conditions, which had no significant impact. This suggests that this computer-tailored intervention might have potential to contribute to the health promotion field.

Authors' Contributions

KDC conceived the study, participated in its design and coordination, analyzed the data, and drafted the manuscript. IDB, GC, and CV conceived the study, participated in its design, and assisted in drafting the manuscript. All authors read, revised, and approved the final manuscript.

KDC is supported by the Research Foundation Flanders (FWO) (postdoctoral research fellowship: FWO11/PDO/097).

Conflicts of Interest

None declared. The authors are the developers of the intervention.

Multimedia Appendix 1

Screenshots of the start to stand tool.

[PDF File (Adobe PDF File), 866KB - jmir_v18i5e96_app1.pdf]

Multimedia Appendix 2

CONSORT-EHEALTH (V 1.6.1) checklist.

[PDF File (Adobe PDF File), 11MB - jmir_v18i5e96_app2.pdf]

References

1. Sedentary Behaviour Research Network. Letter to the editor: standardized use of the terms "sedentary" and "sedentary behaviours". Appl Physiol Nutr Metab 2012 Jun;37(3):540-542. [doi: 10.1139/h2012-024] [Medline: 22540258]
2. Thorp AA, Healy GN, Winkler E, Clark BK, Gardiner PA, Owen N, et al. Prolonged sedentary time and physical activity in workplace and non-work contexts: a cross-sectional study of office, customer service and call centre employees. Int J Behav Nutr Phys Act 2012;9:128 [FREE Full text] [doi: 10.1186/1479-5868-9-128] [Medline: 23101767]
3. Van Acker R, De Meester F. Langdurig zitten: dé uitdaging van de 21ste eeuw: syntheserapport als actuele onderbouw voor de factsheet sedentair gedrag. Brussels: Vlaams Instituut voor Gezondheidspromotie en Ziektepreventie; 2015:1-80.

4. Vandelanotte C, Duncan MJ, Short C, Rockloff M, Ronan K, Happell B, et al. Associations between occupational indicators and total, work-based and leisure-time sitting: a cross-sectional study. BMC Public Health 2013;13:1110 [FREE Full text] [doi: 10.1186/1471-2458-13-1110] [Medline: 24289321]
5. Thorp AA, Owen N, Neuhaus M, Dunstan DW. Sedentary behaviors and subsequent health outcomes in adults a systematic review of longitudinal studies, 1996-2011. Am J Prev Med 2011 Aug;41(2):207-215. [doi: 10.1016/j.amepre.2011.05.004] [Medline: 21767729]
6. Proper KI, Singh AS, van Mechelen W, Chinapaw M. Sedentary behaviors and health outcomes among adults: a systematic review of prospective studies. Am J Prev Med 2011 Feb;40(2):174-182. [doi: 10.1016/j.amepre.2010.10.015] [Medline: 21238866]
7. Healy GN, Wijndaele K, Dunstan DW, Shaw JE, Salmon J, Zimmet PZ, et al. Objectively measured sedentary time, physical activity, and metabolic risk: the Australian Diabetes, Obesity and Lifestyle Study (AusDiab). Diabetes Care 2008 Feb;31(2):369-371. [doi: 10.2337/dc07-1795] [Medline: 18000181]
8. van Uffelen J, Wong J, Chau JY, van der Ploeg H, Riphagen I, Gilson ND, et al. Occupational sitting and health risks: a systematic review. Am J Prev Med 2010 Oct;39(4):379-388. [doi: 10.1016/j.amepre.2010.05.024] [Medline: 20837291]
9. Prince SA, Saunders TJ, Gresty K, Reid RD. A comparison of the effectiveness of physical activity and sedentary behaviour interventions in reducing sedentary time in adults: a systematic review and meta-analysis of controlled trials. Obes Rev 2014 Nov;15(11):905-919 [FREE Full text] [doi: 10.1111/obr.12215] [Medline: 25112481]
10. Shrestha N, Ijaz S, Kukkonen-Harjula K, Kumar S, Nwankwo C. Workplace interventions for reducing sitting at work. Cochrane Database Syst Rev 2015;1:1-132. [doi: 10.1002/14651858.CD010912.pub3] [Medline: 25620219]
11. Tew GA, Posso MC, Arundel CE, McDaid CM. Systematic review: height-adjustable workstations to reduce sedentary behaviour in office-based workers. Occup Med (Lond) 2015 Jul;65(5):357-366. [doi: 10.1093/occmed/kqv044] [Medline: 25934982]
12. Verweij LM, Proper K, Weel A, Hulshof C, van Mechelen W. The application of an occupational health guideline reduces sedentary behaviourincreases fruit intake at work: results from an RCT. Occup Environ Med 2012;69(7):500-507. [doi: 10.1136/oemed-2011-100377] [Medline: 22383591]
13. Stephens SK, Winkler EA, Trost SG, Dunstan DW, Eakin EG, Chastin SF, et al. Intervening to reduce workplace sitting time: how and when do changes to sitting time occur? Br J Sports Med 2014 Jul;48(13):1037-1042. [doi: 10.1136/bjsports-2014-093524] [Medline: 24815544]
14. Neuhaus M, Healy G, Dunstan D, Owen N, Eakin E. Workplace sitting and height-adjustable workstations: A randomized controlled trial. Am J Prev Med 2014;46(1):30-40. [doi: 10.1016/j.amepre.2013.09.009] [Medline: 24355669]
15. Alkhajah TA, Reeves MM, Eakin EG, Winkler EA, Owen N, Healy GN. Sit-stand workstations: a pilot intervention to reduce office sitting time. Am J Prev Med 2012 Sep;43(3):298-303. [doi: 10.1016/j.amepre.2012.05.027] [Medline: 22898123]
16. Pronk NP, Katz AS, Lowry M, Payfer JR. Reducing occupational sitting time and improving worker health: the Take-a-Stand Project, 2011. Prev Chronic Dis 2012;9:E154 [FREE Full text] [Medline: 23057991]
17. Gilson ND, Suppini A, Ryde GC, Brown HE, Brown WJ. Does the use of standing 'hot' desks change sedentary work time in an open plan office? Prev Med 2012 Jan;54(1):65-67. [doi: 10.1016/j.ypmed.2011.10.012] [Medline: 22056630]
18. Carr LJ, Karvinen K, Peavler M, Smith R, Cangelosi K. Multicomponent intervention to reduce daily sedentary time: a randomised controlled trial. BMJ Open 2013;3(10):e003261 [FREE Full text] [doi: 10.1136/bmjopen-2013-003261] [Medline: 24141969]
19. Evans RE, Fawole HO, Sheriff SA, Dall PM, Grant PM, Ryan CG. Point-of-choice prompts to reduce sitting time at work: a randomized trial. Am J Prev Med 2012 Sep;43(3):293-297. [doi: 10.1016/j.amepre.2012.05.010] [Medline: 22898122]
20. Smith L, Pedersen S. The effect of education on compliance to a workplace health and wellbeing intervention: closing the loop. Universal J Public Health 2013;1(3):97-102.
21. Neuhaus M, Eakin EG, Straker L, Owen N, Dunstan DW, Reid N, et al. Reducing occupational sedentary time: a systematic review and meta-analysis of evidence on activity-permissive workstations. Obes Rev 2014 Oct;15(10):822-838. [doi: 10.1111/obr.12201] [Medline: 25040784]
22. De Cocker K, Veldeman C, De Bacquer D, Braeckman L, Owen N, Cardon G, et al. Acceptability and feasibility of potential intervention strategies for influencing sedentary time at work: focus group interviews in executives and employees. Int J Behav Nutr Phys Act Feb 18 2015:12. [doi: 10.1186/s12966-015-0177-5] [Medline: 25881297]
23. Gardner B, Smith L, Lorencatto F, Hamer M, Biddle SJ. How to reduce sitting time? A review of behaviour change strategies used in sedentary behaviour reduction interventions among adults. Health Psychol Rev 2015 Sep 16:1-24. [doi: 10.1080/17437199.2015.1082146] [Medline: 26315814]
24. Neville LM, Milat AJ, O'Hara B. Computer-tailored weight reduction interventions targeting adults: a narrative systematic review. Health Promot J Austr 2009 Apr;20(1):48-57. [Medline: 19402816]
25. Kroeze W, Werkman A, Brug J. A systematic review of randomized trials on the effectiveness of computer-tailored education on physical activity and dietary behaviors. Ann Behav Med 2006 Jun;31(3):205-223. [doi: 10.1207/s15324796abm3103_2] [Medline: 16700634]

26. Lustria ML, Cortese J, Noar SM, Glueckauf RL. Computer-tailored health interventions delivered over the Web: review and analysis of key components. Patient Educ Couns 2009 Feb;74(2):156-173. [doi: 10.1016/j.pec.2008.08.023] [Medline: 18947966]
27. Internet World Stats: Usage and Population Statistics. Bogota, Columbia: Miniwatts Marketing Group; 2014 Nov 30. World Internet Usage and Population Statistics June 30, 2014 - Mid-Year Update URL: http://www.internetworldstats.com/stats.htm [accessed 2015-01-28] [WebCite Cache ID 6Vv7gMvJn]
28. Broekhuizen K, Kroeze W, van Poppel MN, Oenema A, Brug J. A systematicreviewof randomized controlled trials on the effectiveness of computer-tailored physical activity and dietary behavior promotion programs: an update. Ann Behav Med 2012;44(2):259-286. [doi: 10.1007/s12160-012-9384-3] [Medline: 22767052]
29. Kelders SM, Kok RN, Ossebaard HC, Van Gemert-Pijnen J. Persuasive system design does matter: a systematic review of adherence to web-based interventions. J Med Internet Res 2012;14(6):e152 [FREE Full text] [doi: 10.2196/jmir.2104] [Medline: 23151820]
30. De Cocker K, De Bourdeaudhuij I, Cardon G, Vandelanotte C. Theory-driven, web-based, computer-tailored advice to reduce and interrupt sitting at work: development, feasibility and acceptability testing among employees. BMC Public Health 2015;15(1):959 [FREE Full text] [doi: 10.1186/s12889-015-2288-y] [Medline: 26404915]
31. Swartz A, Rote A, Welch W, Maeda H, Hart T, Cho Y, et al. Prompts to disrupt sitting time and increase physical activity at work, 2011-2012. Prev Chronic Dis 2014;11:E73. [doi: _10.5888/pcd11.130318PMCID:] [Medline: 4008949]
32. Deci E, Ryan R. Handbook of Self-determination Research. Rochester, NY: University of Rochester Press; 2002.
33. Ajzen I. The theory of planned behavior. Organ Behav Hum 1991;50(2):179-211.
34. Maes S, Karoly P. Self-regulation assessment and intervention in physical health and illness: a review. Applied Psychol: Int Rev 2005;54(2):267-299.
35. Vandelanotte C, Spathonis KM, Eakin EG, Owen N. Website-delivered physical activity interventions a review of the literature. Am J Prev Med 2007 Jul;33(1):54-64. [doi: 10.1016/j.amepre.2007.02.041] [Medline: 17572313]
36. 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]
37. Rutten GM, Savelberg HH, Biddle SJ, Kremers SP. Interrupting long periods of sitting: good STUFF. Int J Behav Nutr Phys Act 2013;10:1 [FREE Full text] [doi: 10.1186/1479-5868-10-1] [Medline: 23281722]
38. Gollwitzer P. Implementation intentions: strong effects of simple plans. Am Psychol 1999;54:493-503.
39. Chau JY, van der Ploeg HP, Dunn S, Kurko J, Bauman AE. A tool for measuring workers' sitting time by domain: the Workforce Sitting Questionnaire. Br J Sports Med 2011 Dec;45(15):1216-1222. [doi: 10.1136/bjsports-2011-090214] [Medline: 21947817]
40. Clark BK, Thorp AA, Winkler E, Gardiner PA, Healy GN, Owen N, et al. Validity of self-reported measures of workplace sitting time and breaks in sitting time. Med Sci Sports Exerc 2011 Oct;43(10):1907-1912. [doi: 10.1249/MSS.0b013e31821820a2] [Medline: 21926535]
41. Craig CL, Marshall AL, Sjöström M, Bauman AE, Booth ML, Ainsworth BE, et al. International physical activity questionnaire: 12-country reliability and validity. Med Sci Sports Exerc 2003 Aug;35(8):1381-1395. [doi: 10.1249/01.MSS.0000078924.61453.FB] [Medline: 12900694]
42. Vandelanotte C, De Bourdeaudhuij I, Sallis J, Philippaerts R, Sjöström M. Reliability and validity of a computerised and Dutch version of the International Physical Activity Questionnaire (IPAQ). J Phys Act Health 2005;2:63-75.
43. Kozey-Keadle S, Libertine A, Lyden K, Staudenmayer J, Freedson PS. Validation of wearable monitors for assessing sedentary behavior. Med Sci Sports Exerc 2011 Aug;43(8):1561-1567. [doi: 10.1249/MSS.0b013e31820ce174] [Medline: 21233777]
44. Chastin SF, Granat MH. Methods for objective measure, quantification and analysis of sedentary behaviour and inactivity. Gait Posture 2010 Jan;31(1):82-86. [doi: 10.1016/j.gaitpost.2009.09.002] [Medline: 19854651]
45. Crutzen R, Roosjen JL, Poelman J. Using Google Analytics as a process evaluation method for Internet-delivered interventions: an example on sexual health. Health Promot Int 2013 Mar;28(1):36-42 [FREE Full text] [doi: 10.1093/heapro/das008] [Medline: 22377974]
46. Guidelines for Data Processing and Analysis of the International Physical Activity Questionnaire (IPAQ)—Short and Long Forms. 2005 Nov. URL: https://docs.google.com/viewer?a=v&pid=sites&srcid=ZGVmYXVsdGRvbWFpbnx0aGVpcGFxfGd4OjE0NDgxMDk3NDU1YWRlZTM [accessed 2016-04-19] [WebCite Cache ID 6Vv7gMvJn]
47. Cerin E. Statistical approaches to testing the relationships of the built environment with resident-level physical activity behaviorhealth outcomes in cross-sectional studies with cluster sampling. J Plann Lit 2010 Dec 15:151-167. [doi: 10.1177/0885412210386229]
48. Spittaels H, De Bourdeaudhuij I, 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]
49. De Bourdeaudhuij I, Stevens V, Vandelanotte C, Brug J. Evaluation of an interactive computer-tailored nutrition intervention in a real-life setting. Ann Behav Med 2007 Feb;33(1):39-48. [doi: 10.1207/s15324796abm3301_5] [Medline: 17291169]

50. Chau JY, Daley M, Dunn S, Srinivasan A, Do A, Bauman A, van der Ploeg HP. The effectiveness of sit-stand workstations for changing office workers' sitting time: results from the Stand@Work randomized controlled trial pilot. Int J Behav Nutr Phys Act 2014:127. [doi: 10.1186/s12966-014-0127-7] [Medline: 25291960]
51. Matton L, Wijndaele K, Duvigneaud N, Duquet W, Philippaerts R, Thomis M, et al. Reliability and validity of the Flemish Physical Activity Computerized Questionnaire in adults. Res Q Exerc Sport 2007 Sep;78(4):293-306. [doi: 10.1080/02701367.2007.10599427] [Medline: 17941534]
52. Carson V, Wong SL, Winkler E, Healy GN, Colley RC, Tremblay MS. Patterns of sedentary time and cardiometabolic risk among Canadian adults. Prev Med 2014 Aug;65:23-27. [doi: 10.1016/j.ypmed.2014.04.005] [Medline: 24732719]
53. Martin A, Fitzsimons C, Jepson R, Saunders DH, van der Ploeg HP, Teixeira PJ, EuroFIT Consortium. Interventions with potential to reduce sedentary time in adults: systematic review and meta-analysis. Br J Sports Med 2015 Aug;49(16):1056-1063. [doi: 10.1136/bjsports-2014-094524] [Medline: 25907181]
54. Puig-Ribera A, Bort-Roig J, González-Suárez AM, Martínez-Lemos I, Giné-Garriga M, Fortuño J, et al. Patterns of impact resulting from a 'sit less, move more' web-based program in sedentary office employees. PLoS One 2015;10(4):e0122474 [FREE Full text] [doi: 10.1371/journal.pone.0122474] [Medline: 25830782]
55. Mainsbridge CP, Cooley PD, Fraser SP, Pedersen SJ. The effect of an e-health intervention designed to reduce prolonged occupational sitting on mean arterial pressure. J Occup Environ Med 2014 Nov;56(11):1189-1194 [FREE Full text] [doi: 10.1097/JOM.0000000000000243] [Medline: 25376414]
56. Pedersen SJ, Cooley PD, Mainsbridge C. An e-health intervention designed to increase workday energy expenditure by reducing prolonged occupational sitting habits. Work 2014;49(2):289-295. [doi: 10.3233/WOR-131644] [Medline: 23787256]
57. Chau JY, van der Ploeg HP, Merom D, Chey T, Bauman AE. Cross-sectional associations between occupational and leisure-time sitting, physical activity and obesity in working adults. Prev Med 2012;54(3-4):195-200. [doi: 10.1016/j.ypmed.2011.12.020] [Medline: 22227284]
58. Jans MP, Proper KI, Hildebrandt VH. Sedentary behavior in Dutch workers: differences between occupations and business sectors. Am J Prev Med 2007 Dec;33(6):450-454. [doi: 10.1016/j.amepre.2007.07.033] [Medline: 18022060]
59. Tigbe WW, Lean MEJ, Granat MH. A physically active occupation does not result in compensatory inactivity during out-of-work hours. Prev Med 2011;53(1-2):48-52. [doi: 10.1016/j.ypmed.2011.04.018] [Medline: 21575655]
60. Waters LA, Galichet B, Owen N, Eakin E. Who participates in physical activity intervention trials? J Phys Act Health 2011 Jan;8(1):85-103. [Medline: 21297189]
61. Wang PS, Beck AL, McKenas DK, Meneades LM, Pronk NP, Saylor JS, et al. Effects of efforts to increase response rates on a workplace chronic condition screening survey. Med Care 2002 Sep;40(9):752-760. [doi: 10.1097/01.MLR.0000024609.75399.DC] [Medline: 12218766]

Abbreviations

BMI: body mass index
IPAQ: International Physical Activity Questionnaire
SMART: specific, measurable, attainable, relevant, and time-bound
T0: baseline
T1: 1-month follow-up
T2: 3-month follow-up
WSQ: Workforce Sitting Questionnaire

Edited by A Moorhead; submitted 22.10.15; peer-reviewed by J Torous, S Pedersen; comments to author 27.11.15; revised version received 12.01.16; accepted 21.02.16; published 31.05.16.

Please cite as:
De Cocker K, De Bourdeaudhuij I, Cardon G, Vandelanotte C
The Effectiveness of a Web-Based Computer-Tailored Intervention on Workplace Sitting: A Randomized Controlled Trial
J Med Internet Res 2016;18(5):e96
URL: http://www.jmir.org/2016/5/e96/
doi:10.2196/jmir.5266
PMID:27245789

Original Paper

Devices for Self-Monitoring Sedentary Time or Physical Activity: A Scoping Review

James P Sanders[1,2]; Adam Loveday[1,2]; Natalie Pearson[2]; Charlotte Edwardson[1,3]; Thomas Yates[1,3]; Stuart JH Biddle[1,2,4]; Dale W Esliger[1,2]

[1]Leicester-Loughborough Diet, Lifestyle and Physical Activity Biomedical Research Unit, Loughborough University, Loughborough, United Kingdom

[2]National Centre for Sport and Exercise Medicine, School of Sport, Exercise and Health Sciences, Loughborough University, Loughborough, United Kingdom

[3]Diabetes Research Centre, College of Medicine, Biological Sciences and Psychology, University of Leicester, Leicester, United Kingdom

[4]Active Living and Public Health, Institute of Sport, Exercise & Active Living (ISEAL), Victoria University, Melbourne, Australia

Corresponding Author:
James P Sanders
Leicester-Loughborough Diet, Lifestyle and Physical Activity Biomedical Research Unit
Loughborough University
Sir John Beckwith Building
Loughborough, LE11 3TU
United Kingdom
Phone: 44 7538330734
Fax: 44 7538330734
Email: J.Sanders2@lboro.ac.uk

Abstract

Background: It is well documented that meeting the guideline levels (150 minutes per week) of moderate-to-vigorous physical activity (PA) is protective against chronic disease. Conversely, emerging evidence indicates the deleterious effects of prolonged sitting. Therefore, there is a need to change both behaviors. Self-monitoring of behavior is one of the most robust behavior-change techniques available. The growing number of technologies in the consumer electronics sector provides a unique opportunity for individuals to self-monitor their behavior.

Objective: The aim of this study is to review the characteristics and measurement properties of currently available self-monitoring devices for sedentary time and/or PA.

Methods: To identify technologies, four scientific databases were systematically searched using key terms related to behavior, measurement, and population. Articles published through October 2015 were identified. To identify technologies from the consumer electronic sector, systematic searches of three Internet search engines were also performed through to October 1, 2015.

Results: The initial database searches identified 46 devices and the Internet search engines identified 100 devices yielding a total of 146 technologies. Of these, 64 were further removed because they were currently unavailable for purchase or there was no evidence that they were designed for, had been used in, or could readily be modified for self-monitoring purposes. The remaining 82 technologies were included in this review (73 devices self-monitored PA, 9 devices self-monitored sedentary time). Of the 82 devices included, this review identified no published articles in which these devices were used for the purpose of self-monitoring PA and/or sedentary behavior; however, a number of technologies were found via Internet searches that matched the criteria for self-monitoring and provided immediate feedback on PA (ActiGraph Link, Microsoft Band, and Garmin Vivofit) and sedentary time (activPAL VT, the Lumo Back, and Darma).

Conclusions: There are a large number of devices that self-monitor PA; however, there is a greater need for the development of tools to self-monitor sedentary time. The novelty of these devices means they have yet to be used in behavior change interventions, although the growing field of wearable technology may facilitate this to change.

(J Med Internet Res 2016;18(5):e90) doi:10.2196/jmir.5373

KEYWORDS

sitting time; physical activity; measurement; feedback; activity monitor; scoping review

Introduction

Modern environments and technological advancements have radically altered the way we live our lives [1]. The need to undertake purposeful physical activity (PA) has all but disappeared and sedentary behavior, defined as "any waking behavior in a sitting or reclining posture with an energy expenditure ≤1.5 metabolic equivalent" [2] is the dominant behavior. Low levels of moderate-to-vigorous PA (MVPA) have been consistently associated with the risk of developing chronic diseases, such as type 2 diabetes, cardiovascular disease, and some cancers [3]. In addition, increasing the total level of daily movement, such as the number of steps taken, has been strongly inversely associated with the risk of developing chronic diseases [4,5]. There is also mounting evidence that the amount of time spent sedentary is an important determinant of health status independent of PA levels. For example, Wilmot and colleagues [6] found that when comparing those with the highest levels of sedentary behavior to those with the lowest levels, independent of PA levels, there was a 112%, 147%, 90%, and 49% increase in the relative risk of type 2 diabetes, cardiovascular disease, cardiovascular mortality, and all-cause mortality, respectively. Moreover, how sedentary time and PA are accumulated throughout the day may also be important, with frequent breaks in sedentary behavior associated with a healthier metabolic profile [7]. This has necessitated a paradigm shift that focuses on both the accumulation of MVPA (the traditional focus of lifestyle interventions) and the importance of postural allocation throughout the waking hours.

Over the last decade, there has been a plethora of tools developed to support PA and sedentary behavior change, of which the greatest growth has been seen in self-monitoring tools. Self-monitoring is defined as "a person closely and deliberately monitors their own behavior" [8,9] and "allowing the modification of their behaviors to achieve predetermined goals or outcomes" [10] and has a strong theoretical foundation for behavior change. Self-regulation theory posits that self-monitoring precedes self-evaluation of progress made toward one's goal and as well as preceding self-reinforcement of behavior for progress to be made [9]. Furthermore, Control Theory proposes that self-monitoring of behavior, setting goals, receiving feedback, and reviewing relevant goals with feedback work synergistically and are central to self-management and behavioral control [11,12]. Self-monitoring, therefore, can increase an individual's personal responsibility, promote independence, and individuals can create their own pathways toward goal achievement by taking an active rather than passive role [13]. When included in behavior change interventions, self-monitoring has proven to be an effective behavior change strategy across a variety of behaviors, including smoking, diet, and PA, and it is considered a foundation of lifestyle behavior change interventions [12,14].

Traditionally, self-monitoring of PA and sedentary time occurred via paper-based journal methods [14]; more recently, the pedometer became a popular method of self-monitoring for interventions designed to increase PA. Individuals who used pedometers increased their PA by 26.9% from baseline activity levels [15]. Subsequently, advances in technology have led to a proliferation in the number of bodily worn electronic devices becoming available that go beyond simply measuring and providing feedback on the number of steps per day (eg, Fitbit, Jawbone). Along with PA, electronic devices are also starting to measure sitting time, provide real-time feedback, as well as encouraging interruptions in prolonged sitting. It has been suggested that the use of these electronic approaches to self-monitor might lessen the burden of traditional methods and may improve adherence to self-monitoring resulting in greater achievement toward behavioral goals [16].

This increased availability of electronic self-monitoring devices provides an opportunity for researchers to utilize these novel technologies as an aid for behavior change in PA and sedentary behavior on a large scale. Furthermore, wearable technologies are increasingly integrating health care systems. Recent reports from the National Information Board in a review of the National Health Service in the United Kingdom indicate the need for "citizens" to start playing a more active role in their health care by accessing, entering, and uploading data into their own online medical record. Under these new plans, citizens will be able to access and download their detailed medical records as well as contribute to it with information from their personal wearable technology or biosensors [17,18]. In addition, as more health care providers in the United States move to a value-based care system (ie, "reward points" for positive lifestyle alterations that can be redeemed for discounts on a range of products and/or activities), mobile technologies that promote health and well-being by engaging in important health behaviors (eg, increased MVPA) will continue to grow and have the potential to be an integral piece of future health care systems. In light of this, a review of the current tools used to self-monitor PA and/or sedentary time has the potential to be a valuable resource to researchers, clinicians, health care providers, and the general public.

Therefore, it seems timely to review the characteristics and measurement properties (eg, wear location, integrated sensors, outcomes measured) of currently available self-monitoring devices, both those marketed to consumers and those used in research settings, that have been (or could be) utilized in, or developed for, real-time self-monitoring of sedentary behavior and/or PA.

Methods

Searches

The search strategy was built around three groups of keywords: behavior (ie, PA and sedentary behavior), measurement, and population. A detailed description of the keywords used and method of combination can be found in Multimedia Appendix 1. For the purposes of this study, tools were deemed to measure sedentary time if they could measure the wearer's sitting and/or reclining posture.

Scopus, Medline, Web of Science, and the Institute of Electrical and Electronic Engineers (IEEE) databases were searched using these keywords from the inception of the databases to October 1, 2015. In addition, manual searches of personal files were conducted and reference lists of primary studies were screened.

Internet Search Engines

Because of the rapid release of technology in the consumer electronic area, a grey literature search of relevant websites was conducted for technologies that allow for the self-monitoring of PA and sedentary time but may not have made it into the published research to date. Keywords based on the same groups as the database searches were used to search the Internet engines Google, Bing, and Yahoo. Searches were extracted for later review using a specialized browser plug-in. The first 200 search results from each search engine were extracted for further review; this was a pragmatic approach because it was deemed that results after the first 200 were either not relevant or repetitive. This ensured that the results were unaffected by the changing algorithms of Web search engines. Searches were completed on October 1, 2015.

Study Inclusion and Exclusion Criteria

Two sets of inclusion criteria were developed for research articles and websites. For inclusion in the review, studies were required to (1) include adults aged 18 years or older, (2) be published in English, and (3) describe a device that objectively self-monitors PA, physical inactivity, and/or sedentary time/sitting and can, or has the potential to, provide feedback to the user. Traditionally, there would also be a criteria based around study type; however, in order to obtain the widest variety of devices, this was not included.

For inclusion in the review, only websites from manufacturers were included (ie, blogs or consumer reviews pertaining to technologies of interest were excluded) and devices that had the ability to self-monitor and were available for purchase at the time of the review were included.

Data Extraction

Potentially relevant articles were selected by screening titles, screening abstracts, and if abstracts were not available or did not provide sufficient data, the entire article was sought and screened to determine whether it met the inclusion criteria. Relevant websites were selected by screening webpage titles and screening devices on relevant webpages to determine whether it met the inclusion criteria. Data were extracted on standardized forms developed for this review.

Information on the devices was extracted from articles and cross-referenced with the device manufacturer's information. Validity data on each device was not extracted; instead, articles with relevant validity data, where available [19-38], were referenced in the data table because the authors chose to focus this review on the characteristics of the devices to allow the reader to make a judgment about their efficacy as self-monitoring tools.

A 10% subsample of potentially relevant articles retrieved for full-paper screening were extracted by a second author (AL) to determine interrater agreement. Interrater agreement was high (Cohen's kappa=.81). If any discrepancies arose, these were resolved by discussion between authors.

Self-Monitor Scoring

Each device was designated a self-monitoring code: (1) yes, self-monitors PA (Y_{PA}); (2) yes, self-monitors PA and physical inactivity, such as self-monitoring and feedback on lack of movement (Y_{PI}); and (3) yes, self-monitors sedentary time (Y_{ST}).

The different attributes of the self-monitoring devices were based on Control Theory [11]; specifically, the ability to receive feedback (defined as the provision of informative and actionable insights on the performance of the behavior) and the ability to set goals (defined as agreeing on a goal/target defined in terms of the behavior to be achieved) [8]. Aspects included the different types of feedback (eg, vibratory, auditory, omnipresent in the form of colors or lights, or potentially via push notifications). Also included was the timing of the feedback (ie, immediate or delayed). Other features included the way in which the data were portrayed (eg, numeric data/graphical representation of the data). The platform pervasiveness was also included (ie, number of different devices/operating systems the data could be viewed on). Each of these was broken into the feedback attributes that were available on either the device or the backend platform (defined as the smart device/software that the technology connected to). Other attributes included were goal-setting capability of the device and whether the device or associated software could be customized by the end user via some method, usually an application programming interface or software development kit. Textbox 1 provides a detailed description of each self-monitoring attribute. Each attribute was split into whether the attribute was present on the device itself (denoted with "D") or whether it was present on the backend platform (ie, mobile phone/tablet; denoted with "BP").

Textbox 1. Description of the self-monitoring attributes coded.

Auditory: feedback on behavior provided verbally from device (eg, via Sensoria voice-over feedback regarding ground contact from smartphone/smart MP3)

Vibratory: haptic feedback on predetermined behavioral thresholds provided using vibrations (eg, LumoBack)

Omnipresent: feedback that is visible all the time, usually in the form of a progression bar that changes with advancement toward predetermined goals (eg, Fitbit Flower)

Push notification: the delivery of information regarding behavioral goals from a software app to a computing device without a specific request from the user

Immediate: whether the data/feedback are immediate in its return to the user (eg, LumoBack)

Delayed: whether the data/feedback are delayed in its return to the user (eg, ActiGraph)

Numeric: data are returned in the form of numbers/figures or statistics

Graph: data are returned in the form of graphical representation

Written/textual feedback: data are returned in the form of textual feedback

'Ometer (omnipresent meter): data are returned in the form of a growing or shrinking picture/image based on completion toward a predetermined goal (eg, UbiFit Garden)

Application: what operating system the mobile app can be accessed on for viewing the data/feedback

Software: the operating systems it can be accessed on if a piece of computing software is present for use at viewing the data

Website: can the data/feedback be viewed on a website?

Goal-setting capability: can predetermined goals be set by the user?

Customization: can the device or mobile app be customized by the end user (eg, via a software development kit)?

Each device was given a score between 1 and 6 for each attribute of behavior change. This score was used to describe two factors: (1) whether or not that device contained that behavior change attribute and (2) to what extent it did or did not contain the attribute. The self-monitoring scoring system that was used for each attribute was (1) yes; (2) yes, difficulties (eg, proximity to computer); (3) yes, lack of evidence to suggest this; (4) no, but present in future iterations; (5) no, but possible (with application programming interface or software development kit); and (6) not described/featured.

This scoring system was meant to be a descriptive tally of the behavior change attributes and not a judgment on the effectiveness of the various features.

Results

Review Statistics

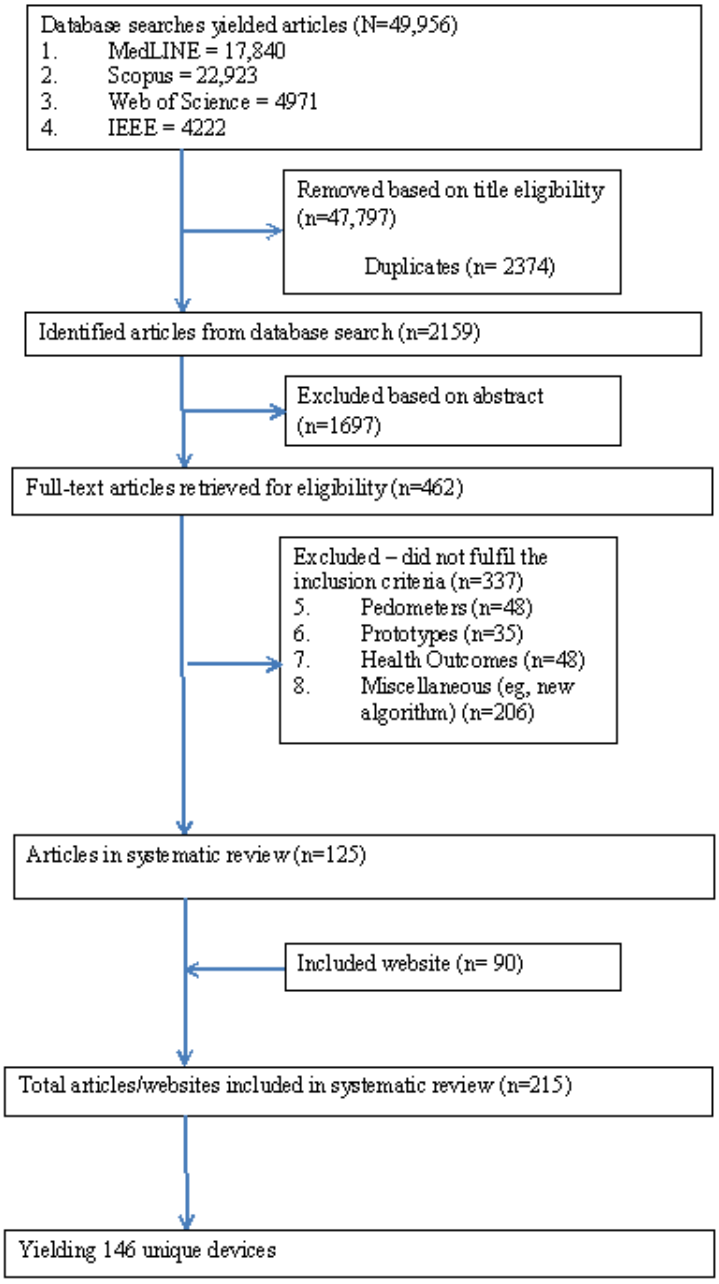
Database searches identified 49,956 articles (Figure 1), of which 462 were deemed to be potentially relevant and were retrieved for full-text analysis. Articles were excluded for a number of reasons (n=337):

1. Pedometer studies: these were excluded if no evidence could be found that the pedometer in question provided temporally stamped data.
2. Prototypes: if the device was not commercially available or if no data currently existed for the prototype and only proof of concept information was available were excluded.
3. Health outcome: articles were excluded if they examined the relationship between behavior (eg, sedentary behavior and/or PA) and a particular health outcome (eg, blood pressure, lipid profile) and the measurement tool of choice was not the main focus of the article.
4. Miscellaneous: articles were excluded if the purpose of the study was to examine a new algorithm or data processing procedure for device analysis.

The remaining 125 studies (on 46 devices) and 90 websites yielded 146 devices (see Multimedia Appendix 2) that were selected for detailed scrutiny. Of these, 64 were further removed because there was no evidence that they were designed for, had been used in, or could readily be modified for real-time self-monitoring purposes or that they were not currently available for purchase.

The remaining 82 [39-119] technologies were included in this review. Of these, 73 [39-110] technologies measured / self-monitored PA, of which 16 [43,45,55,56,58,62-66,81,86,90,91,94,103,107-109] provided some measure of physical inactivity (see Multimedia Appendix 3). In all, 9 [111-119] technologies measured self-monitored sedentary time (Multimedia Appendix 4), 8 [111,112,114-119] of which measured both PA and sedentary time.

Figure 1. Study/website selection.



Physical Activity Self-Monitoring Technologies

Figure 2 displays the number of self-monitoring attributes apparent in each of the devices found to measure/self-monitor PA. The device with the highest number of feedback attributes was the Microsoft Band [77] with 18 of 28 feedback possibilities that were coded. The most common feedback attribute used in the devices found was joint numeric and graphical data feedback on the associated backend platform, with 94% of the devices that self-monitored PA displaying these attributes. The least common form of feedback attribute was auditory feedback from the device (D_Auditory). This particular type of feedback was only present in 2% of cases (Figure 3).

Figure 2. Technologies found that can be used to self-monitor and provide feedback on PA ordered by number of self-monitoring attributes that were found to be present in the technologies.

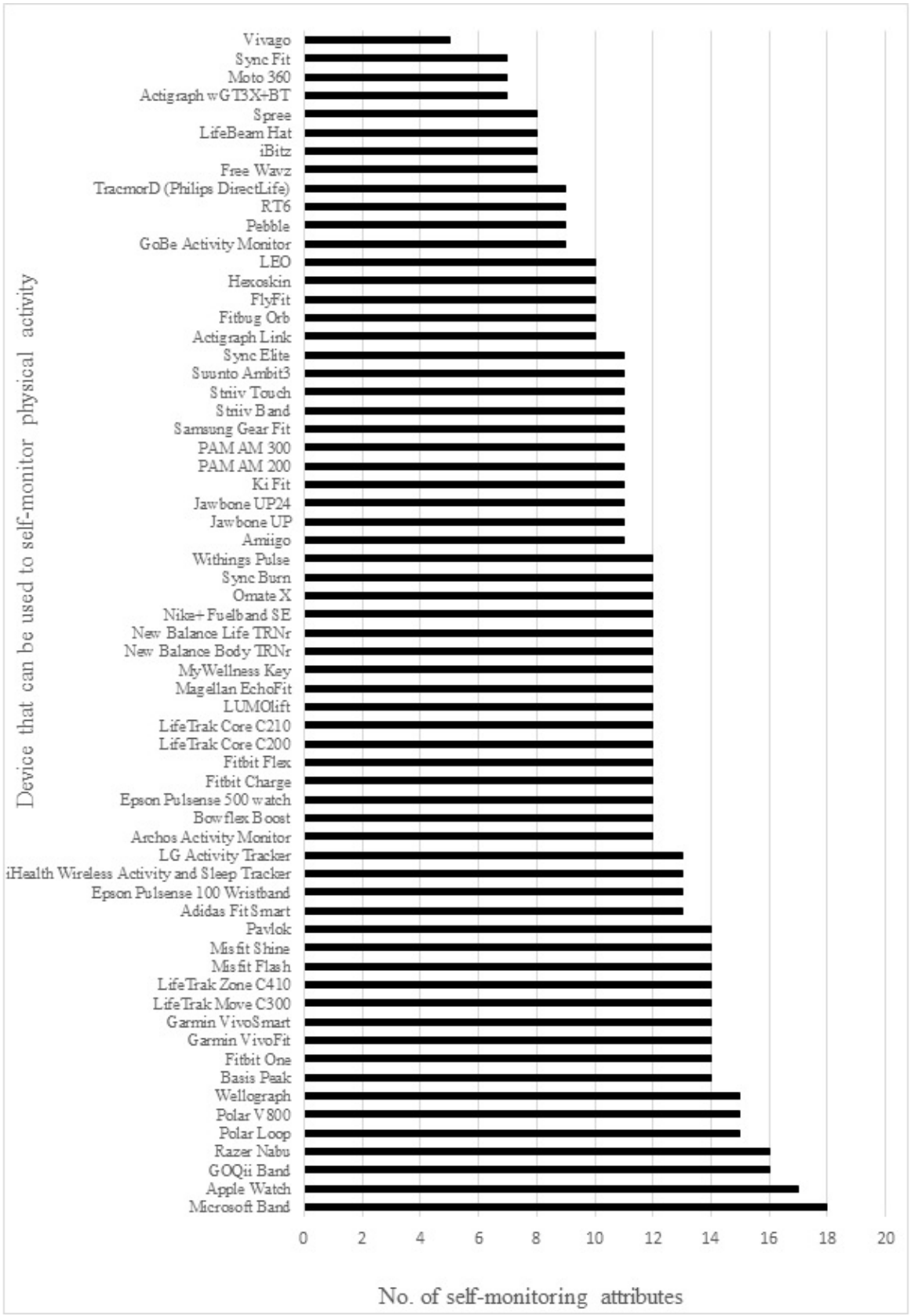
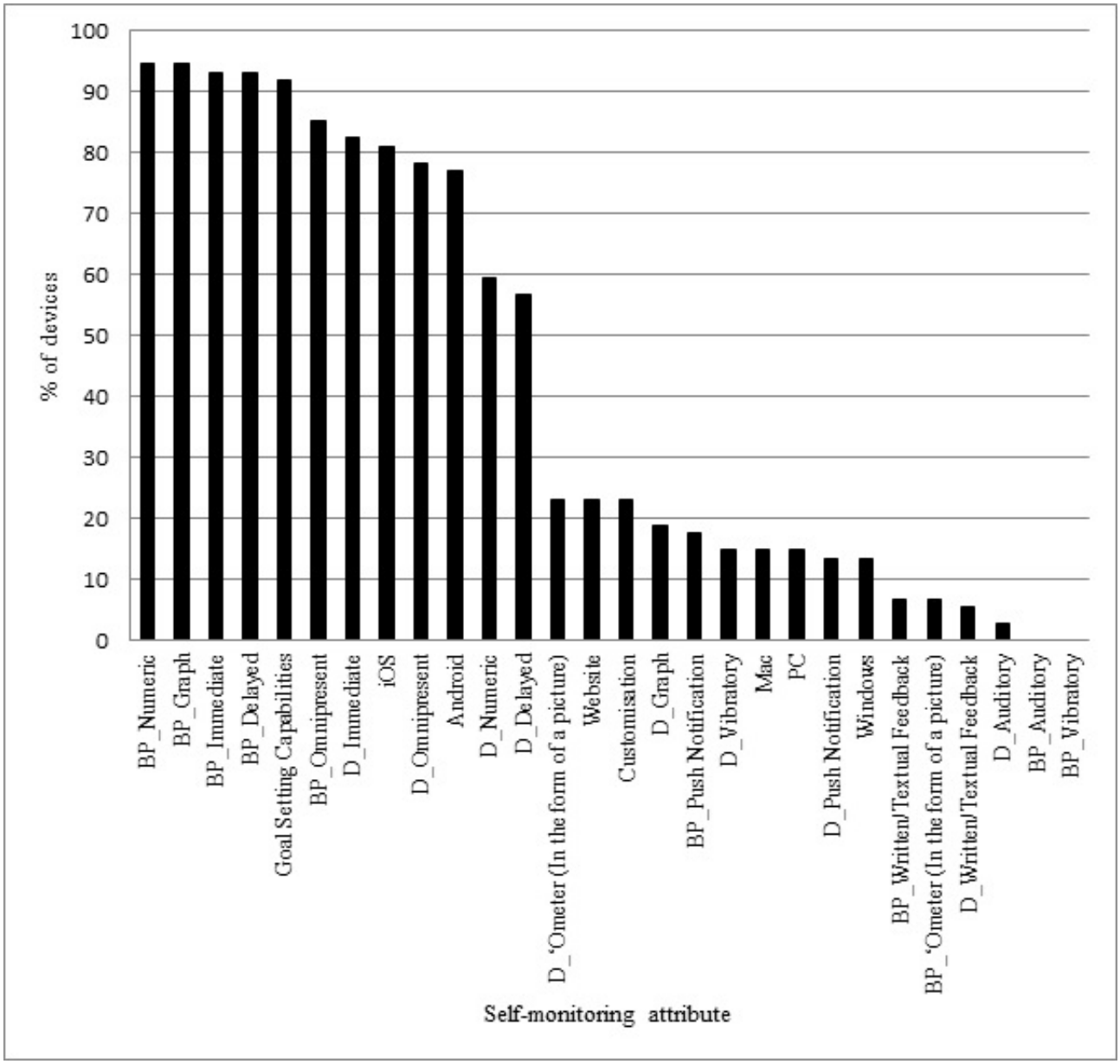


Figure 3. The proportion of devices that could be used to self-monitor and provided feedback on physical activity with specific self-monitoring attributes.



Sedentary Time Self-Monitoring Technologies

Figure 4 displays the number of self-monitoring attributes apparent in each of the devices found to measure/self-monitor sedentary time. Figure 5 documents the popularity of the self-monitoring attributes with sedentary time self-monitoring devices. The device with the highest number of feedback attributes was the Lumo Back posture sensor and feedback coach [17] with 13 of 28 feedback possibilities that were coded. The most common feedback attribute used in the devices found was joint numeric and graphical data feedback on the associated backend platform, with 81% of the devices that self-monitor sedentary time displaying these attributes. The least common form of feedback attribute was push notification of feedback from the device of sedentary time on the device. This particular type of feedback was not present in any of the devices found.

Figure 4. Technologies found that could be used to self-monitor and provided feedback on sedentary time ordered by number of feedback elements in the technologies.

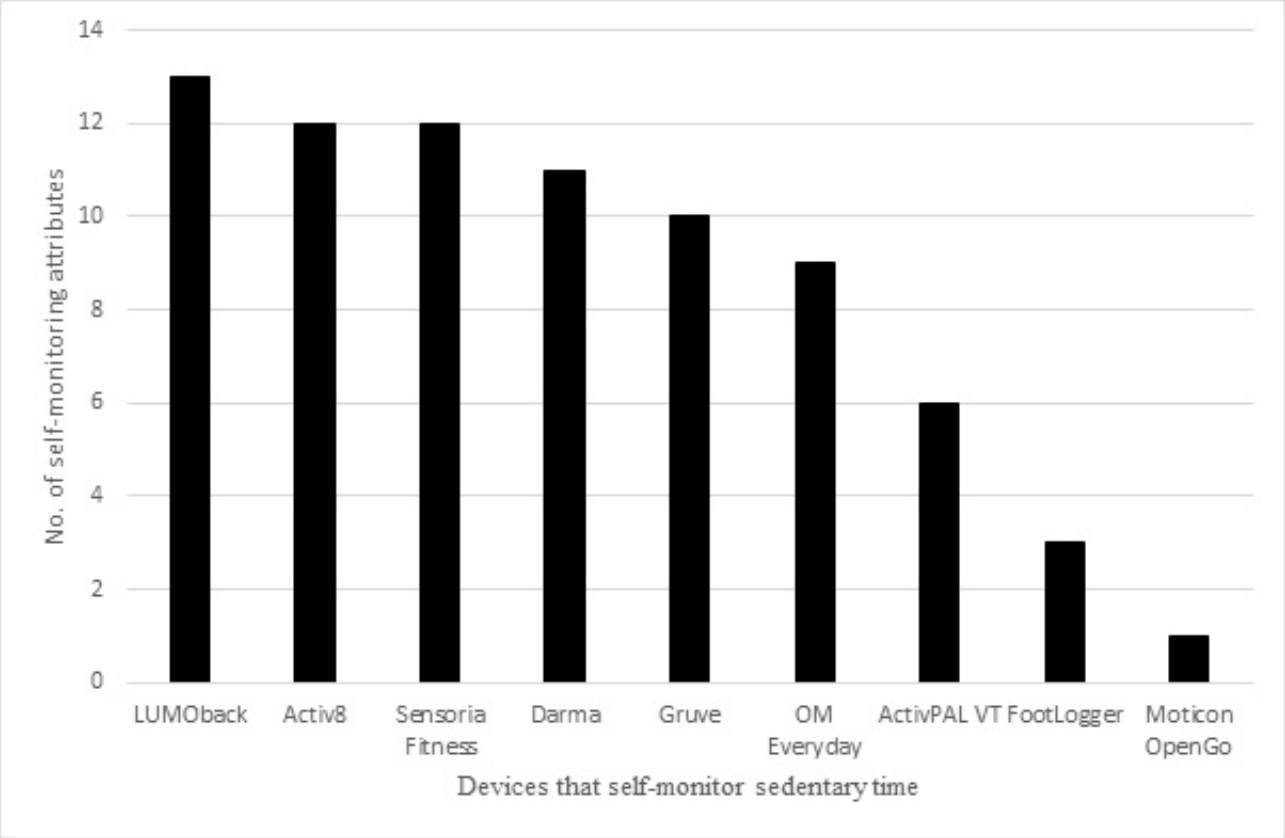
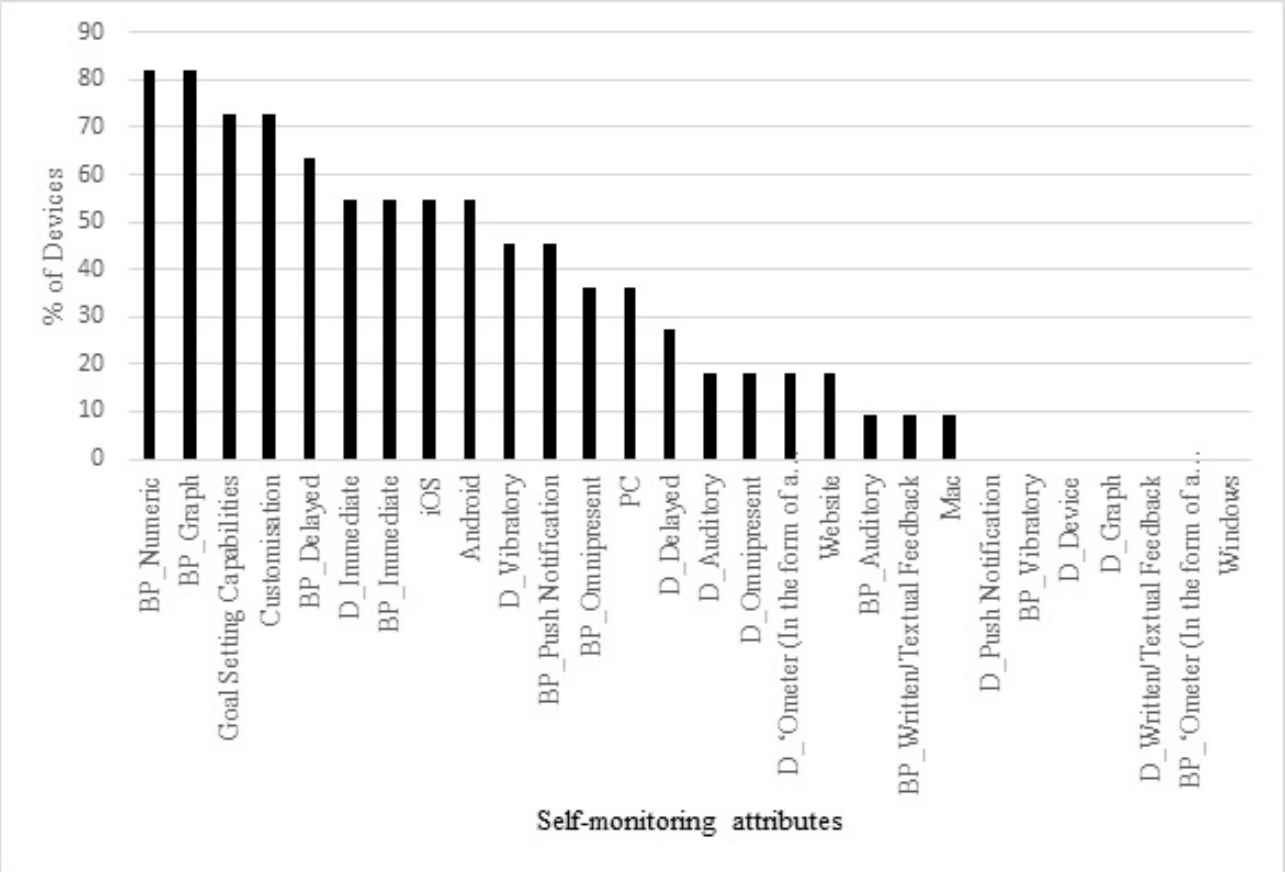


Figure 5. The proportion of devices that could be used to self-monitor and provided feedback on sedentary time with specific self-monitoring attributes.



Discussion

The present systematic review sought to identify current measurement technologies available that could be used for real-time self-monitoring of sedentary time and/or PA. The review identified 125 articles on 46 device and 90 websites, for a combined total of 146 technologies that monitor sedentary time and/or PA. Of these, 82 devices were considered capable of self-monitoring sedentary time and/or PA. These devices can be used by researchers, clinicians, and the general public.

Technologies that self-monitor PA mainly come from the consumer health and fitness market. In general, these devices consist of an accelerometer for activity measurement (steps, calories burned, distance traveled) with varying secondary sensors, including gyroscope, inclinometer, lux sensors, skin sweat sensors, and other sensors that provide additional pieces of information. However, these devices will provide feedback only on PA and increases in PA do not automatically lead to decreases in sedentary time [120]. Additionally, more and more of these devices are providing feedback on not only the amount of PA, but also the length of time spent inactive.

There are devices from both the commercial and research sectors that self-monitor sedentary behavior. These devices tend to measure sedentary time in two different ways. Firstly, posture sensors measure sedentary time either through an accelerometer in conjunction with gravitational components and proprietary algorithms (eg, activPAL) or through the alignment of the area of the body surrounding the pelvic area (ie, pelvic alignment is different depending on standing, sitting, and lying). The other way technologies tend to measure sedentary time is via pressure sensors. These pressure sensors are either located in a sock, shoe, or chair. When placed in a sock or shoe, the pressure can determine standing when there is pressure on the sensor and when there is less pressure the wearer is sitting or lying. Located on a chair, there is a simple binary outcome: when the pressure sensor is active the user is sitting and when it is inactive there is no sitting behavior at that site.

Both these types of devices usually provide feedback either via vibratory feedback (eg, Jawbone UP) or via an omnipresent display on the device (eg, Garmin Vivofit). These devices tend to, but not exclusively, connect to a mobile app for feedback on the PA and sedentary time. For PA, this usually takes the form of energy expenditure or proprietary company points (eg, Nike Fuel). For sedentary time, this usually takes the form of time spent sitting (eg, LumoBack) These mobile apps allow the wearer to receive real-time continuous feedback along with goal-setting capabilities and customization of type and timing of feedback; this is an aspect not traditionally offered by research devices.

With the plethora of devices now available (see Figure 6 for an example of popular devices), with differing attributes and cost, it is unsurprising that these devices are growing in popularity. However, and perhaps paradoxically, there are a small number of devices specifically designed to measure sitting time. Furthermore, the small number of devices that do provide feedback on sitting were not either originally designed for its measurement (eg, LumoBack) or are still primarily research tools to be used in scientific study (eg, ActivPAL VT).

Self-monitoring technologies need to provide real-time feedback on aspects of PA and sitting that are personalized and relevant to the individual (ie, the attributes of real-time feedback must resonate with the individual and not be simply information that has been presupposed for them). Additionally, the immediate feedback should be of a low cognitive load so that it can resonate immediately with the end user [121,122]. For example, the Fitbit one has a growing flower as a feedback indication of progression toward a user-defined goal. Using a pictorial representation of this nature will resonate easier with the user [123,124]. However, more detailed information on the temporal patterning of the behavior, for example, should be accessible from a mobile app, website, or software. Furthermore, the likelihood of the feedback being acted upon could be increased if it is provided in a manner that is context aware. In other words, the feedback must be given at a time when it can be acted upon by the user. For example, to reduce sitting, it should provide feedback while watching television rather than sitting in an exam or during a prolonged dental procedure. If these attributes could be integrated into a single device, it would help facilitate its use by differing populations regardless of technological ability. These devices need to have a substantial battery life and memory capacity at a reasonable cost. For this to occur, there is a need for cooperative work across different research disciplines and commercial fields to develop these context-aware, personalized feedback devices.

Not every user will have the same needs and the presentation of actionable information will need to be tailored to fit individual needs. In addition, simply providing more medical data to patients not only fails to guarantee improved outcomes, but also could potentially lead to negative consequences [125]. Activity trackers have poor evidence of prolonged use, with a conservatively estimated one-third discontinuing use by 6 months after initiation [126]. A recent study of several tools to encourage medication adherence in older adults, a major area of focus of mHealth developers, found that the most common descriptors participants used to describe their experience with the devices were “frustrating” and “challenging” [127]. In another study of the usage of a dietary app to promote healthy eating, investigators found that fewer than 3% used the app for at least 1 week and fewer than 10% of these individuals made positive changes in their diet [128]. Users require consumer-friendly devices and apps that are self-reinforcing and enjoyable to use. These goals might be accomplished with the use of incentives, gamification, and social networks to promote managed competition/cooperation among peers or family members.

In order for the promise of wearable technology to be fully realized, consumers, providers, and health care systems must be able to trust the reliability, privacy, and security of their data as well as the devices that collect and share it. Although regulatory oversight is often considered to be an impediment to the rapid propagation of innovative technologies, the existence of potential scams that could harm the end user dictates the need for some level of oversight. Globally, there is a great deal of uncertainty around wearable technology regulation; there are

numerous countries that have no regulatory framework, whereas the other countries that do have a framework are still in their infancy and being actively refined [129,130].

Wearable technology users are also concerned about the privacy and ownership of their health data. In the era of big data, it is critical that the terms of ownership of personal data, most especially medical data, be unmistakably stated—not buried in the commonly unread and then accepted terms of use agreements—with users required to explicitly consent whenever their data are sold or transmitted to others [131].

One of the benefits of mHealth is easier accessibility to pertinent health care data, but this increased availability to both consumers and providers creates the potential for substantial security risks. Because of the small size of the device, it becomes easier to inadvertently lose or easier to steal, which may mean that the information stored on the device becomes accessible to others.

As consumer demand for wearable sensor increases, health care providers will face the possibility of being inundated by a flood of patient data. This will create a number of difficult challenges, including the potential requirement for 24/7 oversight, the need to summarize multiparameter, continuously collected data into a usable and clinically meaningful format [132].

The strengths of this review are the systematic approach taken and the comprehensive range of technologies found. However, there are some limitations. Due to the nature of articles included, it was not possible to present data on the validity and reliability of the devices in their ability to measure sedentary time. Similarly, due to the fact that self-monitoring using objective measurement tools is in its infancy, there are gaps in the literature as to whether these devices truly work as self-monitors; consequently, we cannot comment on how useful or valid they are in these settings. However, validity data are important. Users of self-monitoring technologies must be able to trust in the feedback that is being returned to them otherwise they may become disenfranchised with the tool and the behavior change tool. Therefore, incorporating important valid data with the feedback tools means additional value can be added to the consumers and potentially more potent behavior change.

In conclusion, the authors believe that this review is the first of its kind to systematically describe the wide breadth of devices that self-monitor and provide feedback on PA and sedentary behavior. There has been an explosion in the number of devices that measure PA and there is a greater need for the development of tools that specifically measure sitting time. Cooperative work between engineers, computer scientists, and academics in relevant fields is needed to develop these technologies that provide real-time, personalized, context-aware feedback to aid in the reduction in sitting time and its detrimental effect on cardiometabolic health independent of PA. This could potentially lead to the use of these devices in a health care setting as part of the increasing value-based care systems that are starting to arise in the United States or as a diagnostic tool, which is beginning to be implemented in the National Health Service in the United Kingdom.

This scoping review provides a record of a plethora of devices with information on their capabilities both in terms of their ability to measure behavior and to provide feedback to the user, providing a foundation for clinical, research, and public health use. Future studies are needed to further investigate the validity of these devices and their feasibility in increasing PA and/or decreasing sedentary time and the public health impact this may produce.

Figure 6. Example of devices discussed in this review. Clockwise from top left: Fitbit Charge HR, Garmin Vivofit, Jawbone Up, Nike Fuelband SE, Lumo Back Posture Sensor, and a mobile app.

Acknowledgments

TY and CLE acknowledge support from the National Institute for Health Research Collaboration for Leadership in Applied Health Research and Care-East Midlands (NIHR CLAHRC-EM) and the Leicester Clinical Trials Unit, as well as the NIHR Leicester-Loughborough Diet, Lifestyle and Physical Activity Biomedical Research Unit, which is a partnership between University Hospitals of Leicester NHS Trust, Loughborough University, and the University of Leicester. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR, or the Department of Health.

This work was supported in part by a Medical Research Council grant in LLHW Promoting Physical Activity in Older Age 2012 (MRC Reference: MR/K025090/1). This work was additionally supported by an EPSRC Small Equipment Grant titled: Upgrading the Small Equipment base for Early Career Researchers in the Engineering and Physical Sciences (EP/K03118X/1).

Conflicts of Interest

None declared.

Multimedia Appendix 1

Search Strategy.

[PDF File (Adobe PDF File), 40KB - jmir_v18i5e90_app1.pdf]

Multimedia Appendix 2

Supplementary Table 1 - All Devices.

[XLSX File (Microsoft Excel File), 58KB - jmir_v18i5e90_app2.xlsx]

Multimedia Appendix 3

Supplementary Table 2 - Devices that self monitor physical activity.

[XLSX File (Microsoft Excel File), 45KB - jmir_v18i5e90_app3.xlsx]

Multimedia Appendix 4

Supplementary Table 3 - Devices that Self-monitor Sedentary behaviour.

[XLSX File (Microsoft Excel File), 19KB - jmir_v18i5e90_app4.xlsx]

References

1. Katzmarzyk PT, Mason C. The physical activity transition. J Phys Act Health 2009 May;6(3):269-280. [Medline: 19564654]
2. Sedentary Behaviour Research Network. Letter to the editor: standardized use of the terms "sedentary" and "sedentary behaviours". Appl Physiol Nutr Metab 2012 Jun;37(3):540-542. [doi: 10.1139/h2012-024] [Medline: 22540258]
3. World Health Organization. Global Recommendations on Physical Activity for Health. Geneva: World Health Organization; 2010.
4. Ogilvie D, Foster CE, Rothnie H, Cavill N, Hamilton V, Fitzsimons CF, Scottish Physical Activity Research Collaboration. Interventions to promote walking: systematic review. BMJ 2007 Jun 9;334(7605):1204 [FREE Full text] [doi: 10.1136/bmj.39198.722720.BE] [Medline: 17540909]
5. Morris JN, Hardman AE. Walking to health. Sports Med 1997 May;23(5):306-332. [Medline: 9181668]
6. Wilmot EG, Edwardson CL, Achana FA, Davies MJ, Gorely T, Gray LJ, et al. Sedentary time in adults and the association with diabetes, cardiovascular disease and death: systematic review and meta-analysis. Diabetologia 2012 Nov;55(11):2895-2905. [doi: 10.1007/s00125-012-2677-z] [Medline: 22890825]
7. Healy GN, Dunstan DW, Salmon J, Cerin E, Shaw JE, Zimmet PZ, et al. Breaks in sedentary time: beneficial associations with metabolic risk. Diabetes Care 2008 Apr;31(4):661-666. [doi: 10.2337/dc07-2046] [Medline: 18252901]
8. Michie S, Richardson M, Johnston M, Abraham C, Francis J, Hardeman W, et al. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. Ann Behav Med 2013 Aug;46(1):81-95. [doi: 10.1007/s12160-013-9486-6] [Medline: 23512568]
9. Kanfer F, Gaelick-Buys L. Self-management methods. In: Help People Change: A Textbook of Methods. 4th ed. Elmsford, NY: Pergamon Press; 1991:309-355.
10. Fogg B. Persuasive Technology: Using Computers to Change What We Think and Do. San Francisco, CA: Morgan Kaufmann; 2003.
11. Carver CS, Scheier MF. Control theory: a useful conceptual framework for personality-social, clinical, and health psychology. Psychol Bull 1982 Jul;92(1):111-135. [Medline: 7134324]

12. Michie S, Abraham C, Whittington C, McAteer J, Gupta S. Effective techniques in healthy eating and physical activity interventions: a meta-regression. Health Psychol 2009 Nov;28(6):690-701. [doi: 10.1037/a0016136] [Medline: 19916637]
13. Schachter SC. Quality of life for patients with epilepsy is determined by more than seizure control: the role of psychosocial factors. Expert Rev Neurother 2006 Jan;6(1):111-118. [doi: 10.1586/14737175.6.1.111] [Medline: 16466318]
14. Burke LE, Wang J, Sevick MA. Self-monitoring in weight loss: a systematic review of the literature. J Am Diet Assoc 2011 Jan;111(1):92-102 [FREE Full text] [doi: 10.1016/j.jada.2010.10.008] [Medline: 21185970]
15. Bravata DM, Smith-Spangler C, Sundaram V, Gienger AL, Lin N, Lewis R, et al. Using pedometers to increase physical activity and improve health: a systematic review. JAMA 2007 Nov 21;298(19):2296-2304. [doi: 10.1001/jama.298.19.2296] [Medline: 18029834]
16. Turner-McGrievy GM, Beets MW, Moore JB, Kaczynski AT, Barr-Anderson DJ, Tate DF. Comparison of traditional versus mobile app self-monitoring of physical activity and dietary intake among overweight adults participating in an mHealth weight loss program. J Am Med Inform Assoc 2013 May 1;20(3):513-518 [FREE Full text] [doi: 10.1136/amiajnl-2012-001510] [Medline: 23429637]
17. National Information Board. Personalised Health and Care 2020: Using Data and Technlogy to Transform Outcomes for Patients and Citizens - A Framework for Action. 2014. URL: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/384650/NIB_Report.pdf [accessed 2016-04-04] [WebCite Cache ID 6gWCAwAIP]
18. National Information Board. GOV.UK. 2015. Delivering the five year forward view: personalised health and care 2020 URL: https://www.gov.uk/government/publications/implementing-personalised-health-and-care-2020/delivering-the-five-year-forward-view [accessed 2016-04-04] [WebCite Cache ID 6gWCIcEIE]
19. Dannecker K, Petro S, Melanson EL, Browning RC. Accuracy of Fitbit activity monitor to predict energy expenditure with and without classification of activities. Med Sci Sport Exer 2011;43(Suppl 1):62. [doi: 10.1249/01.MSS.0000402857.55729.ab]
20. Dannecker KL, Sazonova NA, Melanson EL, Sazonov ES, Browning RC. A comparison of energy expenditure estimation of several physical activity monitors. Med Sci Sports Exerc 2013 Nov;45(11):2105-2112 [FREE Full text] [doi: 10.1249/MSS.0b013e318299d2eb] [Medline: 23669877]
21. Bonomi AG, Plasqui G, Goris AH, Westerterp KR. Estimation of free-living energy expenditure using a novel activity monitor designed to minimize obtrusiveness. Obesity (Silver Spring) 2010 Sep;18(9):1845-1851. [doi: 10.1038/oby.2010.34] [Medline: 20186133]
22. Plasqui G, Bonomi AG, Westerterp KR. Daily physical activity assessment with accelerometers: new insights and validation studies. Obes Rev 2013 Jun;14(6):451-462. [doi: 10.1111/obr.12021] [Medline: 23398786]
23. Bergamin M, Ermolao A, Sieverdes JC, Zaccaria M, Zanuso S. Validation of the mywellness key in walking and running speeds. J Sports Sci Med 2012;11(1):57-63 [FREE Full text] [Medline: 24149122]
24. Sieverdes JC, Wickel EE, Hand GA, Bergamin M, Moran RR, Blair SN. Reliability and validity of the Mywellness Key physical activity monitor. Clin Epidemiol 2013;5:13-20 [FREE Full text] [doi: 10.2147/CLEP.S38370] [Medline: 23378783]
25. Herrmann SD, Hart TL, Lee CD, Ainsworth BE. Evaluation of the MyWellness Key accelerometer. Br J Sports Med 2011 Feb;45(2):109-113. [doi: 10.1136/bjsm.2009.062182] [Medline: 19736173]
26. Ryan CG, Grant PM, Tigbe WW, Granat MH. The validity and reliability of a novel activity monitor as a measure of walking. Br J Sports Med 2006 Sep;40(9):779-784 [FREE Full text] [doi: 10.1136/bjsm.2006.027276] [Medline: 16825270]
27. Grant PM, Ryan CG, Tigbe WW, Granat MH. The validation of a novel activity monitor in the measurement of posture and motion during everyday activities. Br J Sports Med 2006 Dec;40(12):992-997 [FREE Full text] [doi: 10.1136/bjsm.2006.030262] [Medline: 16980531]
28. Kozey-Keadle S, Libertine A, Lyden K, Staudenmayer J, Freedson PS. Validation of wearable monitors for assessing sedentary behavior. Med Sci Sports Exerc 2011 Aug;43(8):1561-1567. [doi: 10.1249/MSS.0b013e31820ce174] [Medline: 21233777]
29. Chastin SF, Granat MH. Methods for objective measure, quantification and analysis of sedentary behaviour and inactivity. Gait Posture 2010 Jan;31(1):82-86. [doi: 10.1016/j.gaitpost.2009.09.002] [Medline: 19854651]
30. Vanhelst J, Hurdiel R, Mikulovic J, Bui-Xuân G, Fardy P, Theunynck D, et al. Validation of the Vivago wrist-worn accelerometer in the assessment of physical activity. BMC Public Health 2012;12:690 [FREE Full text] [doi: 10.1186/1471-2458-12-690] [Medline: 22913286]
31. Takacs J, Pollock CL, Guenther JR, Bahar M, Napier C, Hunt MA. Validation of the Fitbit One activity monitor device during treadmill walking. J Sci Med Sport 2014 Sep;17(5):496-500. [doi: 10.1016/j.jsams.2013.10.241] [Medline: 24268570]
32. Lee J, Kim Y, Welk GJ. Validity of consumer-based physical activity monitors. Med Sci Sports Exerc 2014 Sep;46(9):1840-1848. [doi: 10.1249/MSS.0000000000000287] [Medline: 24777201]
33. Harrington DM, Welk GJ, Donnelly AE. Validation of MET estimates and step measurement using the ActivPAL physical activity logger. J Sports Sci 2011 Mar;29(6):627-633. [doi: 10.1080/02640414.2010.549499] [Medline: 21360402]
34. Rosenberger ME, Buman MP, Haskell WL, Mcconnell MV, Carstensen LL. Twenty-four hours of sleep, sedentary behavior, and physical activity with nine wearable devices. Med Sci Sport Exer 2016;48(3):457-465. [doi: 10.1249/MSS.0000000000000778]

35. Ferguson T, Rowlands AV, Olds T, Maher C. The validity of consumer-level, activity monitors in healthy adults worn in free-living conditions: a cross-sectional study. Int J Behav Nutr Phys Act 2015;12:42 [FREE Full text] [doi: 10.1186/s12966-015-0201-9] [Medline: 25890168]
36. Tucker WJ, Bhammar DM, Sawyer BJ, Buman MP, Gaesser GA. Validity and reliability of Nike + Fuelband for estimating physical activity energy expenditure. BMC Sports Sci Med Rehabil 2015;7:14 [FREE Full text] [doi: 10.1186/s13102-015-0008-7] [Medline: 26751385]
37. Kooiman TJ, Dontje ML, Sprenger SR, Krijnen WP, van der Schans CP, de Groot M. Reliability and validity of ten consumer activity trackers. BMC Sports Sci Med Rehabil 2015;7:24 [FREE Full text] [doi: 10.1186/s13102-015-0018-5] [Medline: 26464801]
38. Bai Y, Welk GJ, Nam YH, Lee JA, Lee J, Kim Y, et al. Comparison of consumer and research monitors under semistructured settings. Med Sci Sports Exerc 2015 Jul 6:151-158. [doi: 10.1249/MSS.0000000000000727] [Medline: 26154336]
39. ActiGraph LLC. 2015. ActiGraph Link URL: http://www.actigraphcorp.com/ [accessed 2015-10-26] [WebCite Cache ID 6cZAImFqx]
40. ActiGraph LLC. 2013. ActiGraph wGT3X-BT Monitor URL: http://www.actigraphcorp.com/ [accessed 2015-10-26] [WebCite Cache ID 6cZAImFqx]
41. Adidas International Trading BV. 2014. Adidas MiCoach - Fit Smart URL: http://micoach.adidas.com/fitsmart/ [accessed 2015-10-27] [WebCite Cache ID 6caSmu04u]
42. Amiigo. 2014. Amiigo Wristband URL: https://amiigo.com/ [accessed 2015-10-27] [WebCite Cache ID 6caSyBWfX]
43. Apple Inc. 2014. Apple Watch URL: http://www.apple.com/watch/ [accessed 2015-10-27] [WebCite Cache ID 6caT8N3Ep]
44. ARCHOS. 2014. ARCHOS Activity Tracker URL: http://www.archos.com/gb/products/objects/cself/aat/index.html [accessed 2015-10-27] [WebCite Cache ID 6caTAIdri]
45. BASIS Science Inc. 2014. Basis Peak URL: http://www.mybasis.com/ [accessed 2015-10-27] [WebCite Cache ID 6caTFWjEt]
46. Nautilus Inc. 2013. Bowflex Boost URL: http://www.bowflexboost.com/bowflex-boost-us/homepage.jsp [accessed 2015-10-27] [WebCite Cache ID 6caZU956w]
47. Epson America Inc. 2014. Epson Pulsense - Fitness Watch & Band URL: http://www.epson.com/cgi-bin/Store/jsp/Landing/pulsense-fitness-sensing-watches.do?ref=van_pulsense [accessed 2015-10-27] [WebCite Cache ID 6caTWdEFa]
48. Fitbit Inc. 2014. Fitbit Charge URL: https://www.fitbit.com/uk [accessed 2015-10-27] [WebCite Cache ID 6caTSdmPl]
49. Fitbit Inc. 2014. Fitbit Flex URL: https://www.fitbit.com/uk [accessed 2015-10-27] [WebCite Cache ID 6caTSdmPl]
50. Fitbit Inc. 2012. Fitbit One URL: https://www.fitbit.com/uk [accessed 2015-10-27] [WebCite Cache ID 6caTSdmPl]
51. Fitbit Inc. 2014. Fitbit Surge URL: http://www.fitbit.com/uk/home [accessed 2015-10-28] [WebCite Cache ID 6cceuNs9o]
52. Fitbug Ltd. 2014. Fitbug Orb URL: https://www.fitbug.com/g/orb?lng=en_UK [accessed 2015-10-27] [WebCite Cache ID 6caTb9Mxo]
53. FlyFit. 2014. FlyFit URL: http://www.myflyfit.com/ [accessed 2015-10-27] [WebCite Cache ID 6caTeZEP6]
54. Free Wavz. 2014. Free Wavz -wire free smart earphones URL: http://www.freewavz.com/ [accessed 2015-10-27] [WebCite Cache ID 6caThI16u]
55. Garmin Ltd. 2014. Garmin VivoFit URL: https://buy.garmin.com/en-GB/GB/watches-wearable-technology/wearables/vivofit-/prod143405.html [accessed 2015-10-27] [WebCite Cache ID 6caTzNYgO]
56. Garmin Ltd. 2014. Garmin VivoSmart URL: https://buy.garmin.com/en-GB/GB/watches-wearable-technology/wearables/vivosmart-/prod154886.html [accessed 2015-10-27] [WebCite Cache ID 6caTx1U8K]
57. HealBe Corporation. 2014. HealBe GoBe Activity Monitor URL: http://healbe.com/us/ [accessed 2015-10-27] [WebCite Cache ID 6caUUPZ3j]
58. GOQii Inc. 2014. GOQii Life URL: http://goqii.com/ [accessed 2015-10-27] [WebCite Cache ID 6caULu8nl]
59. Hexoskin. 2013. Hexoskin - wearable body metrics URL: http://www.hexoskin.com/ [accessed 2015-10-27] [WebCite Cache ID 6caV09LI8]
60. GeoPalz LLC. 2014. iBitz URL: http://ibitz.com/ [accessed 2015-10-27] [WebCite Cache ID 6caU2lYtg]
61. iHealth Lab Inc. 2014. iHealth Wireless Activity and Sleep Tracker URL: http://www.ihealthlabs.com/fitness-devices/wireless-activity-and-sleep-tracker/ [accessed 2015-10-27] [WebCite Cache ID 6caW4tgH6]
62. Jawbone. 2014. Jawbone UP URL: https://jawbone.com/ [accessed 2015-10-27] [WebCite Cache ID 6caXzadLa]
63. Jawbone. 2014. Jawbone UP2 URL: https://jawbone.com/ [accessed 2015-10-28] [WebCite Cache ID 6ccfAHlvj]
64. Jawbone. 2014. Jawbone UP24 URL: https://jawbone.com/ [accessed 2015-10-27] [WebCite Cache ID 6caXzadLa]
65. Jawbone. 2014. Jawbone UP3 URL: https://jawbone.com/ [accessed 2015-10-28] [WebCite Cache ID 6ccfAHlvj]
66. Jawbone. 2015. Jawbone UP4 URL: https://jawbone.com/ [accessed 2015-10-28] [WebCite Cache ID 6ccfAHlvj]
67. Ki Performance. 2014. Ki Fit URL: http://www.kiperformance.co.uk/ [accessed 2015-10-27] [WebCite Cache ID 6caY2S3dy]
68. Gesture Logic. 2014. LEO URL: http://leohelps.com/ [accessed 2016-04-04] [WebCite Cache ID 6gWE5brqi]
69. LG Electronics. 2014. LG Lifeband Touch Activity Tracker URL: http://www.lg.com/us/mobile-accessories/lg-FB84-BL-activity-tracker [accessed 2015-10-27] [WebCite Cache ID 6caYF5IKF]
70. LifeBeam Technologies Ltd. 2014. LifeBeam Hat URL: http://life-beam.com/shop/smart-hat/ [accessed 2015-10-27] [WebCite Cache ID 6caYv579j]

71. Salutron Inc. 2014. LifeTrak Move C200 URL: http://lifetrakusa.com/product/core-c200/ [accessed 2015-10-27] [WebCite Cache ID 6caaJrnBO]
72. Salutron Inc. 2014. LifeTrak Core C210 URL: https://lifetrakusa.com/product/move-c300/ [accessed 2015-10-27] [WebCite Cache ID 6caaLC9cP]
73. Salutron Inc. 2014. LifeTrak Move C300 URL: https://lifetrakusa.com/product/move-c300/ [accessed 2015-10-27] [WebCite Cache ID 6caaLC9cP]
74. Salutron Inc. 2014. LifeTrak Zone C410 URL: http://lifetrakusa.com/product/zone-c410/ [accessed 2015-10-27] [WebCite Cache ID 6caaTrPHX]
75. LumoBodyTech Inc. 2014. LumoLift URL: http://www.lumobodytech.com/ [accessed 2015-10-27] [WebCite Cache ID 6caYyxIOq]
76. MiTAC International Corporation. 2014. Magellan EchoFit Series URL: http://www.magellangps.com/lp/echofit/index.html [accessed 2015-10-27] [WebCite Cache ID 6caZ6uTnb]
77. Microsoft. 2014. Microsoft Band URL: http://www.microsoft.com/microsoft-band/en-gb [accessed 2015-10-27] [WebCite Cache ID 6caZ2VZTD]
78. Misfit Wearables. 2014. Misfit Flash Fitness + Sleep Monitor URL: http://misfit.com/ [accessed 2015-10-27] [WebCite Cache ID 6caZ4owci]
79. Misfit Wearables. 2014. Misfit Shine Fitness + Sleep Tracker URL: http://misfit.com/ [accessed 2015-10-27] [WebCite Cache ID 6caZ4owci]
80. Motorola Mobility. 2014. Moto 360 URL: http://www.motorola.co.uk/Moto-360/moto-360-gb.html [accessed 2015-10-27] [WebCite Cache ID 6caZGLYJr]
81. Technogym. 2011. MyWellness Key URL: http://www.technogym.com/gb/products.html [accessed 2015-10-27] [WebCite Cache ID 6cacUjhV1]
82. New Balance. 2014. New Balance Life TRNr URL: http://www.nike.com/gb/en_gb/c/nikeplus-fuel [accessed 2015-10-27] [WebCite Cache ID 6caZYgcOn]
83. New Balance. 2014. New Balance Body TRNr URL: http://www.newbalance.com/pd/body-trnr/52569.html [accessed 2015-10-27] [WebCite Cache ID 6caZWAr15]
84. Nike Inc. 2013. Nike+ Fuelband SE URL: http://www.nike.com/gb/en_gb/c/nikeplus-fuelband [accessed 2016-04-04] [WebCite Cache ID 6gWF1uiBk]
85. Omate. 2014. Omate X URL: http://www.omate.com/index.html [accessed 2015-10-27] [WebCite Cache ID 6caZaAiHs]
86. PAM bv. 2008. PAM AM200 URL: http://www.pam.com/index.php?pid=1 [accessed 2015-10-27] [WebCite Cache ID 6caZhMUBz]
87. PAM bv. 2009. PAM AM300 URL: http://www.pam.com/index.php?pid=1 [accessed 2015-10-27] [WebCite Cache ID 6caZhMUBz]
88. Behavioral Technology Group. 2014. Pavlok URL: http://pavlok.com/ [accessed 2015-10-27] [WebCite Cache ID 6caTMAbln]
89. Pebble. 2012. Pebble Smart Watch URL: https://getpebble.com/ [accessed 2016-04-04] [WebCite Cache ID 6gWF7C3eP]
90. Polar Electro. 2014. Polar Loop URL: http://www.polar.com/en/products/fitness_crosstraining/loop2 [accessed 2015-10-28] [WebCite Cache ID 6ccAUfmeU]
91. Polar Electro. 2014. Polar V800 Sports Watch URL: http://www.polar.com/en/products/maximize_performance/running_multisport/V800 [accessed 2015-04-01] [WebCite Cache ID 6caZyVDzb]
92. Razer Inc. 2014. Razer Nabu URL: http://www.razerzone.com/gb-en/nabu [accessed 2016-04-04] [WebCite Cache ID 6gWFADKzW]
93. Razer Inc. 2015. Razr Nabu X URL: http://www.razerzone.com/nabu-x [accessed 2016-04-04] [WebCite Cache ID 6gWFDin78]
94. Stayhealthy Research. 2013. RT6 URL: http://www.rt6tracker.com/ [accessed 2015-10-27] [WebCite Cache ID 6cac0X9or]
95. SAMSUNG. 2014. SAMSUNG Gear Fit URL: http://www.samsung.com/us/mobile/wearable-tech/SM-R3500ZKAXAR [accessed 2015-10-27] [WebCite Cache ID 6caaVl02g]
96. Spree Wearables. 2014. Spree URL: http://spreewearables.com/ [accessed 2015-10-27] [WebCite Cache ID 6cabu3diL]
97. Stayhealthy Research. 2014. SAM URL: http://www.rt6tracker.com/ [accessed 2015-10-28] [WebCite Cache ID 6ccf2zj4t]
98. Striiv Inc. 2014. Striiv Band URL: http://www.striiv.com/striivband/ [accessed 2016-04-04] [WebCite Cache ID 6gWFJvswR]
99. Striiv Inc. 2015. Striiv Fusion URL: https://jawbone.com/ [accessed 2015-10-28] [WebCite Cache ID 6ccfAHlvj]
100. Striiv Inc. 2015. Striiv Fusion Bio URL: https://jawbone.com/ [accessed 2015-10-28] [WebCite Cache ID 6ccfAHlvj]
101. Striiv Inc. 2015. Striiv Fusion Lite URL: https://jawbone.com/ [accessed 2015-10-28] [WebCite Cache ID 6ccfAHlvj]
102. Striiv Inc. 2014. Striiv Touch URL: http://www.striiv.com/striivtouch/ [accessed 2016-04-04] [WebCite Cache ID 6gWFMPlvy]
103. Amer Sports Brand. 2014. Suunto Ambit3 URL: http://www.suunto.com/en-GB/Products/sports-watches/Suunto-Ambit3-Sport/Suunto-Ambit3-Sport-Blue/ [accessed 2015-10-27] [WebCite Cache ID 6caStKgJf]
104. SYNC. 2014. SYNC Burn URL: http://syncactive.com/products/activity-trackers/ [accessed 2015-10-27] [WebCite Cache ID 6cacQAaPF]

105. SYNC. 2014. SYNC Elite URL: http://syncactive.com/products/fitness/ [accessed 2015-10-27] [WebCite Cache ID 6cacRhDP8]
106. SYNC. 2014. SYNC Fit URL: http://syncactive.com/products/fitness/ [accessed 2015-10-27] [WebCite Cache ID 6cacOIDMi]
107. Philips. 2009. Direct Life URL: http://www.directlife.philips.com/ [accessed 2015-04-01] [WebCite Cache ID 6caZws1a7]
108. Vivago SAS. 2012. Vivago URL: http://www.vivagowellness.fr/ [accessed 2015-10-27] [WebCite Cache ID 6cacYtgoL]
109. Wellograph Company Ltd. 2013. Wellograph URL: https://wellograph.com/ [accessed 2015-10-27] [WebCite Cache ID 6caccVhVs]
110. Withings. 2012. Withings Pulse URL: http://www.withings.com/uk/en/store/details/pulse-O2 [accessed 2015-04-01] [WebCite Cache ID 6ccB89QB7]
111. VitaMove. 2013. Activ8 URL: http://www.vitamove.nl/products/activ8/ [accessed 2015-10-27] [WebCite Cache ID 6cb5jVvYx]
112. PAL technologies. 2014. ActivPAL VT URL: http://www.paltechnologies.com/products/ [accessed 2015-10-27] [WebCite Cache ID 6cb5t224A]
113. Darma Inc. 2014. Darma URL: http://darma.co/ [accessed 2015-10-27] [WebCite Cache ID 6cb60PhjR]
114. 3L Labs. 2014. Footlogger URL: http://www.footlogger.com/ [accessed 2015-10-27] [WebCite Cache ID 6cb68zetk]
115. Gruve Technologies. 2012. Gruve URL: http://www.gruvetechnologies.com/ [accessed 2015-10-27] [WebCite Cache ID 6cb6I8ULz]
116. LumoBodyTech Inc. 2012. Lumo Back URL: http://www.lumobodytech.com/lumo-back/ [accessed 2015-10-27] [WebCite Cache ID 6cb6S7fhj]
117. Moticon. 2013. Moticon OpenGo URL: http://www.moticon.de/products/science-research [accessed 2015-10-27] [WebCite Cache ID 6cb6YODda]
118. OM Everyday. 2013. OM Signal URL: http://www.omsignal.com/ [accessed 2015-10-27] [WebCite Cache ID 6cb6lzVUT]
119. Heapsylon. 2013. Sensoria Fitness Socks URL: http://www.sensoriafitness.com/ [accessed 2015-10-27] [WebCite Cache ID 6cb6wGcSf]
120. Prince SA, Saunders TJ, Gresty K, Reid RD. A comparison of the effectiveness of physical activity and sedentary behaviour interventions in reducing sedentary time in adults: a systematic review and meta-analysis of controlled trials. Obes Rev 2014 Nov;15(11):905-919 [FREE Full text] [doi: 10.1111/obr.12215] [Medline: 25112481]
121. Glendon A, Cernecca L. Young drivers' responses to anti-speeding and anti-drink-driving messages. Transp Res Part F Traffic Psychol Behav 2003 Sep 24;6(3):197-216. [doi: 10.1016/S1369-8478(03)00026-3]
122. Glendon AI, Walker BL. Can anti-speeding messages based on protection motivation theory influence reported speeding intentions? Accid Anal Prev 2013 Aug;57:67-79. [doi: 10.1016/j.aap.2013.04.004] [Medline: 23643708]
123. Sweller J. Cognitive load during problem solving: effects on learning. Cogn Sci 1988;12(2):257-285. [doi: 10.1207/s15516709cog1202_4]
124. Varshney U. A model for improving quality of decisions in mobile health. Decis Support Syst 2014 Jun;62:66-77. [doi: 10.1016/j.dss.2014.03.005]
125. O'Kane MJ, Bunting B, Copeland M, Coates VE. Efficacy of self monitoring of blood glucose in patients with newly diagnosed type 2 diabetes (ESMON study): randomised controlled trial. BMJ 2008 May 24;336(7654):1174-1177 [FREE Full text] [doi: 10.1136/bmj.39534.571644.BE] [Medline: 18420662]
126. Ledger D, McCaffrey D. Inside Wearables: How the Science of Human Behavior Change Offers the Secret to Long-Term Engagement Internet. Cambridge, MA: Endeavour Partners; 2014. URL: http://endeavourpartners.net/assets/Endeavour-Partners-Wearables-White-Paper-20141.pdf [accessed 2015-12-20] [WebCite Cache ID 6dvVD4Akf]
127. Grindrod KA, Li M, Gates A. Evaluating user perceptions of mobile medication management applications with older adults: a usability study. JMIR Mhealth Uhealth 2014;2(1):e11 [FREE Full text] [doi: 10.2196/mhealth.3048] [Medline: 25099993]
128. Helander E, Kaipainen K, Korhonen I, Wansink B. Factors related to sustained use of a free mobile app for dietary self-monitoring with photography and peer feedback: retrospective cohort study. J Med Internet Res 2014;16(4):e109 [FREE Full text] [doi: 10.2196/jmir.3084] [Medline: 24735567]
129. GOV.UK Medicine & Healthcare products Regulatory Agency. 2014 Aug 08. Medical device stand-alone software including apps URL: https://www.gov.uk/government/publications/medical-devices-software-applications-apps/medical-device-stand-alone-software-including-apps [accessed 2015-12-20] [WebCite Cache ID 6dvSP1DKo]
130. US Food and Drug Administration. 2014. Mobile medical applications URL: http://www.fda.gov/medicaldevices/digitalhealth/mobilemedicalapplications/default.htm#f [WebCite Cache ID 6dvSbJn93]
131. Steinhubl SR, Muse ED, Topol EJ. The emerging field of mobile health. Sci Transl Med 2015 Apr 15;7(283):283rv3 [FREE Full text] [doi: 10.1126/scitranslmed.aaa3487] [Medline: 25877894]
132. Yang YT, Silverman RD. Mobile health applications: the patchwork of legal and liability issues suggests strategies to improve oversight. Health Aff (Millwood) 2014 Feb;33(2):222-227. [doi: 10.1377/hlthaff.2013.0958] [Medline: 24493764]

Abbreviations

PA: physical activity

MVPA: moderate-to-vigorous physical activity

Edited by S Kitsiou; submitted 24.11.15; peer-reviewed by K Lyden, A Miner; comments to author 14.12.15; revised version received 18.01.16; accepted 21.01.16; published 04.05.16.

Please cite as:
Sanders JP, Loveday A, Pearson N, Edwardson C, Yates T, Biddle SJH, Esliger DW
Devices for Self-Monitoring Sedentary Time or Physical Activity: A Scoping Review
J Med Internet Res 2016;18(5):e90
URL: http://www.jmir.org/2016/5/e90/
doi:10.2196/jmir.5373
PMID:27145905

Original Paper

Carbohydrate Estimation by a Mobile Phone-Based System Versus Self-Estimations of Individuals With Type 1 Diabetes Mellitus: A Comparative Study

Daniel Rhyner[1,2*], MD; Hannah Loher[2*], MD; Joachim Dehais[1], MSc; Marios Anthimopoulos[1,3], PhD; Sergey Shevchik[1], PhD; Ransford Henry Botwey[1], MSc; David Duke[4], PhD; Christoph Stettler[2], MD; Peter Diem[2], MD; Stavroula Mougiakakou[1,2], PhD

[1]ARTORG Center for Biomedical Engineering Research, University of Bern, Bern, Switzerland

[2]Division of Endocrinology, Diabetes and Clinical Nutrition, Bern University Hospital – Inselspital, Bern, Switzerland

[3]Department of Emergency Medicine, Bern University Hospital – Inselspital, Bern, Switzerland

[4]Roche Diabetes Care, Inc, Indianapolis, IN, United States

[*]these authors contributed equally

Corresponding Author:
Stavroula Mougiakakou, PhD
ARTORG Center for Biomedical Engineering Research
University of Bern
Murtenstrasse 50
Bern, 3008
Switzerland
Phone: 41 316327592
Fax: 41 316327576
Email: stavroula.mougiakakou@artorg.unibe.ch

Abstract

Background: Diabetes mellitus is spreading throughout the world and diabetic individuals have been shown to often assess their food intake inaccurately; therefore, it is a matter of urgency to develop automated diet assessment tools. The recent availability of mobile phones with enhanced capabilities, together with the advances in computer vision, have permitted the development of image analysis apps for the automated assessment of meals. GoCARB is a mobile phone-based system designed to support individuals with type 1 diabetes during daily carbohydrate estimation. In a typical scenario, the user places a reference card next to the dish and acquires two images using a mobile phone. A series of computer vision modules detect the plate and automatically segment and recognize the different food items, while their 3D shape is reconstructed. Finally, the carbohydrate content is calculated by combining the volume of each food item with the nutritional information provided by the USDA Nutrient Database for Standard Reference.

Objective: The main objective of this study is to assess the accuracy of the GoCARB prototype when used by individuals with type 1 diabetes and to compare it to their own performance in carbohydrate counting. In addition, the user experience and usability of the system is evaluated by questionnaires.

Methods: The study was conducted at the Bern University Hospital, “Inselspital” (Bern, Switzerland) and involved 19 adult volunteers with type 1 diabetes, each participating once. Each study day, a total of six meals of broad diversity were taken from the hospital’s restaurant and presented to the participants. The food items were weighed on a standard balance and the true amount of carbohydrate was calculated from the USDA nutrient database. Participants were asked to count the carbohydrate content of each meal independently and then by using GoCARB. At the end of each session, a questionnaire was completed to assess the user’s experience with GoCARB.

Results: The mean absolute error was 27.89 (SD 38.20) grams of carbohydrate for the estimation of participants, whereas the corresponding value for the GoCARB system was 12.28 (SD 9.56) grams of carbohydrate, which was a significantly better performance (P=.001). In 75.4% (86/114) of the meals, the GoCARB automatic segmentation was successful and 85.1% (291/342) of individual food items were successfully recognized. Most participants found GoCARB easy to use.

Conclusions: This study indicates that the system is able to estimate, on average, the carbohydrate content of meals with higher accuracy than individuals with type 1 diabetes can. The participants thought the app was useful and easy to use. GoCARB seems to be a well-accepted supportive mHealth tool for the assessment of served-on-a-plate meals.

(J Med Internet Res 2016;18(5):e101) doi:10.2196/jmir.5567

KEYWORDS

diabetes mellitus, type 1; carbohydrate counting; computer vision systems; food recognition; meal assessment; mobile phone; food volume estimation

Introduction

The disease burden related to diabetes mellitus (hereafter, diabetes) is high and is still rising globally, fuelled by the global rise in the prevalence of obesity and unhealthy lifestyles. The latest estimates show a global prevalence of 387 million people with diabetes in 2015, which is expected to rise to 592 million by 2035 [1]. Of every US $9 spent in health care, US $1 is spent on diabetes corresponding to an estimated expenditure of US $612 billion worldwide in 2014 [2]. Premature morbidity, mortality, reduced life expectancy, financial, and other costs of diabetes make it one of the key public health conditions of the 21st century. Type 1 and type 2 diabetes are the two main types, with type 2 diabetes accounting for the majority (>85%) of total diabetes prevalence [1,3]. Approximately 5% to 15% of patients suffer from type 1 diabetes. Type 1 diabetes is an autoimmune process that completely destroys the insulin-producing pancreatic beta cells, leaving the individuals dependent on exogenous insulin. Achieving a metabolic state close to the physiological is a very challenging task and involves glucose monitoring, insulin treatment, diet management, controlled physical activity, and continuous education on diabetes management. Despite the availability of new drugs, advanced educational programs, and technical solutions that permit continuous glucose monitoring and subcutaneous insulin infusion (insulin pumps), episodes of hypo- and hyperglycemia are still common and the risk of micro- and macrovascular diseases related to diabetes (eg, cardiovascular diseases, retinopathy, and nephropathy) is high.

Estimating the amount of insulin to deliver is one of most challenging and essential tasks in the everyday life of individuals with diabetes. Counting the carbohydrate content of meals to be consumed is a cornerstone of optimal insulin dose estimation. It has been shown that an inaccuracy of ±10 grams does not impair postprandial glycemic control in children [4], but a variation of ±20 grams significantly affects postprandial glycemia [5]. To this end, individuals with diabetes attend nutritional courses on carbohydrate counting; however, according to numerous studies, patients face recurrent difficulties [6-9]. Meals on a plate are especially prone to carbohydrate underestimation, which underlines the emergent need for novel approaches to carbohydrate estimation. The debate on how to optimally estimate carbohydrate intake and facilitate advanced carbohydrate-counting regimes is ongoing and controversial [10].

Over the last two decades, a number of apps have been introduced to track carbohydrate consumption. A recent systematic review analyzed 31 digital approaches to record food intake and nutrition [11]. These systems use different devices—mostly mobile phones, followed by personal computers, and, in older studies, personal digital assistants. They are primarily designed for users who are overweight or obese, with diabetes mellitus, or who want to stay healthy. However, the vast majority of the apps reviewed, except for two using a barcode scanning function, relied on manual input of data either by typing or by selecting a food type from a database. Other proposed systems employ trained health care workers located at a remote location and giving advice on food type, volume, or calorie content. One example is the remote food photography method that relies on the user to take mobile phone pictures of her/his meals, snacks, and beverages, which are then remotely transferred to a trained person who compares the meal to an existing image database and rates its nutrient or caloric content [12]. However, these methods have many limitations in terms of user-friendliness, cost-effectiveness, availability on a large scale, and reproducibility. In a recent study [13], a mobile phone app using augmented reality was proposed that facilitates the estimation of a meal's carbohydrate content by considering its three-dimensional (3D) shape as drawn by the user on the mobile phone screen. According to this study, the system helped users improve their carbohydrate estimation skills, although the required manual input is still burdensome.

All methods presented for nonpacked food assessment have either relied on manual user input or involve a remotely located human component. The recent advances in computer vision and the widespread use of mobile phones with enhanced capabilities have permitted the development of image analysis apps for the automatic assessment of food intake. The input of such an app is a number of images or a short video of the upcoming meal as captured by the user's mobile phone camera. A series of image analysis steps follows, executed either on the mobile phone or on a remote server, to recognize the types and quantities of the meal's food items and to estimate the corresponding nutritional information. Over the last few years, several systems have been proposed with different assumptions, input requirements, and algorithmic approaches [14-16].

The first attempt was made by DiaWear, a system aiming to provide calorie information to diabetic patients [17]. DiaWear considers four fast-food classes and requires one image, with the foods placed on a lighter background and separated from one another. The nutritional content is found directly from reference tables without attempting to estimate the volume. The Pittsburgh Fast-food Image Dataset was created to test food recognition algorithms and contains seven classes of fast-food products [18,19] but, again, the food portion is not considered.

The system developed within the Technology-Assisted Dietary Assessment project uses one meal image and covers 19 food classes [20]. Food volume is estimated by fitting spherical or prismatic 3D models on the detected food areas. However, the use of a single image to estimate the food's volume inevitably depends on weak assumptions about its 3D shape. The use of multiple images by some systems enhanced the results, especially for the volume estimation. The Food Intake Visual and Voice Recognizer [21] uses a short video of the meal from which three images are extracted and used for dense stereo reconstruction and volume estimation. For recognition, 26 food types are supported; nevertheless, the system relies on the user to enter food types through speech because this enhances classification accuracy. DietCam [14], a system for calorie estimation, requires three images or a video of a meal and segments/recognizes food items by visually matching them to a database. The volume is estimated by sparsely reconstructing its 3D shape and fitting an appropriate model. However, critical system characteristics, such as the food classes considered, are not specified. Pouladzadeh et al [15] proposed a system for calorie measurement considering 30 food classes and using the thumb as a reference object. The first image is taken from the top to estimate the food area, the second from the side to estimate height, and then the two values are multiplied to obtain the volume. However, capturing the dish from the side causes occlusions, while assuming constant height for all the food items will introduce large errors as does any misplacement of the thumb.

GoCARB is a novel system for carbohydrate estimation designed for individuals with type 1 diabetes and aims to achieve carbohydrate estimation with an error less than 20 grams per meal. In a previous study, the system was technically evaluated using 24 multifood dishes [16]. The entire evaluation was performed by the researchers involved under controlled conditions. The results showed that the prototype was able to estimate the carbohydrate content with a mean absolute error in the order of 10% (SD 13) or mean 6 (SD 8) carbohydrate grams per meal. The scope of the present study is to assess the performance of the GoCARB prototype when used by individuals with type 1 diabetes and compare it to their own performance in carbohydrate counting.

Methods

GoCARB System

GoCARB is a novel system that aims to support individuals with type 1 diabetes in carbohydrate counting. The system runs on Android mobile phones and uses computer vision to estimate the carbohydrate content of meals. An overview of the system's main intermediate results is presented in Figure 1, whereas the corresponding flowchart is shown in Figure 2. For each estimation, the user places a reference card next to the meal and acquires two images using the camera of the mobile phone. The images are acquired at 0±3 degrees and at 15±3 degrees from the vertical axis crossing the center of the dish, respectively. A graphical user interface supports the user in choosing the optimal angles for each image based on the built-in motion sensors of the mobile phone; the user is only allowed to take the picture when the frame turns green (Figure 3 a-b). The system was designed to have minimum assumptions; namely, that the scene contains only one dish, which should be round, and that there should be no occlusion among the different food items in the dish. After acquisition, the images are transmitted to a dedicated server via Wi-Fi or the mobile network, where a series of computer vision operations are performed. All computer vision modules run on the server, whereas the mobile phone is used for image acquisition, calculation of carbohydrate values, and visualization of the results.

The first step in the series is to detect the dish and automatically segment the different food items in it [22] (Figure 3 c). The dish is detected by extracting the edges of the image and applying a robust fitting paradigm to find an elliptical plate border. The accuracy of this step was estimated to be over 99%, which is essential because the rest of the system relies on it. Automatic segmentation then grows homogenous color regions on a grid inside the dish and merges them on the basis of their color distance and mutual edge size until a minimum size is reached. The accuracy of this module is over 88% and, if it fails, an interactive segmentation tool can be used, which is reliable even for the most difficult cases. To this end, the user has to roughly indicate the position of each food item by touching the screen of the phone (Figure 3 d). These user-given points are used as "seeds" instead of a grid to grow homogenous color regions that correspond to each food item. As soon as the segmentation result is approved by the user, automatic recognition [23] is applied to each segmented food item by using color and texture features fed to a support vector machine. Nine broad food classes are considered; namely, pasta, potatoes, meat, breaded (eg, schnitzel), rice, green salad/vegetables, mashed potatoes, carrots, and beans. The accuracy of this automatic recognition is over 85%. If the result is wrong, the user may correct the system by choosing the right food class from an ordered list in accordance with the confidence of the classifier (Figure 3 e). The food's 3D shape is then reconstructed [24], utilizing both acquired images and the reference card. Key points are detected and matched between the two images to define the orientation and location of the images in space. Using this information, all image pixels are put in correspondence between the two images and their disparity provides the depth used to build the 3D model. By using the 3D model and the segmentation results, the volume of each item is calculated and used to obtain the corresponding carbohydrate content based on the US Department of Agriculture (USDA) nutrient database [25]. For each of the nine food classes, all relevant entries in the database are identified and the mean carbohydrate density is assigned to it. Finally, the results are transmitted back to the mobile phone and displayed to the user (Figure 3 f). More detailed information on technical specifications can be found elsewhere [16,22-24].

Figure 1. Overview of the GoCARB system.

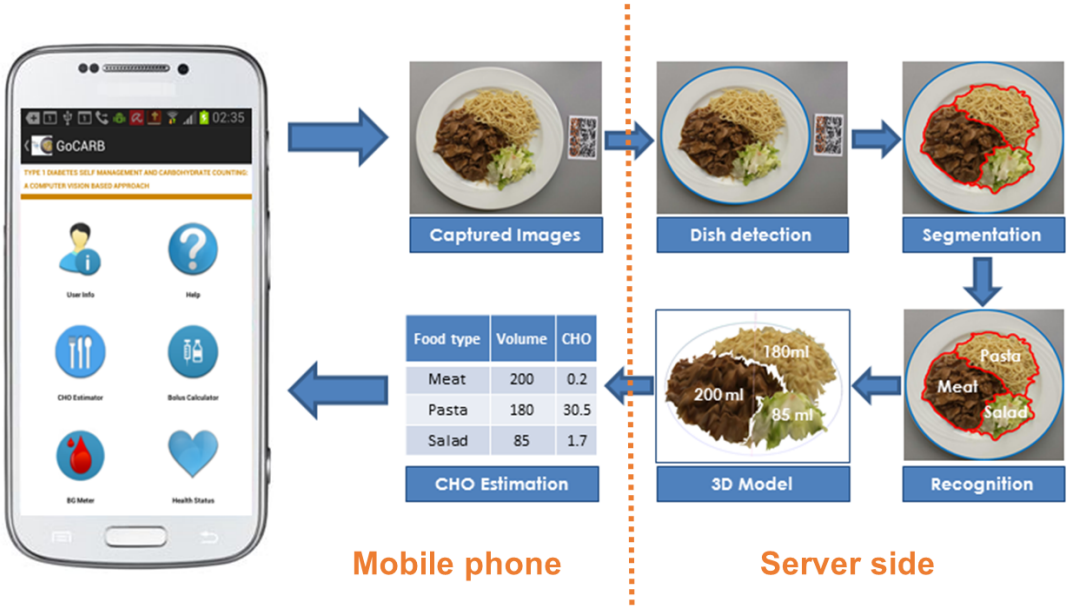


Figure 2. Flowchart of the GoCARB system.

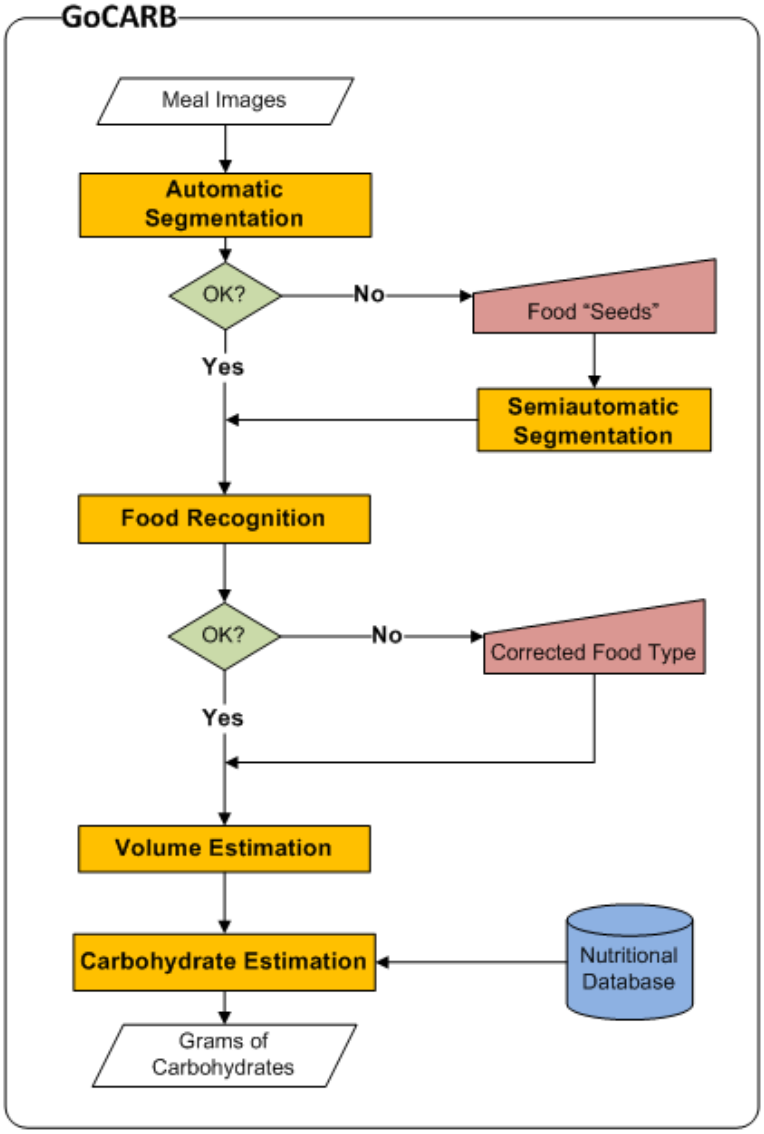
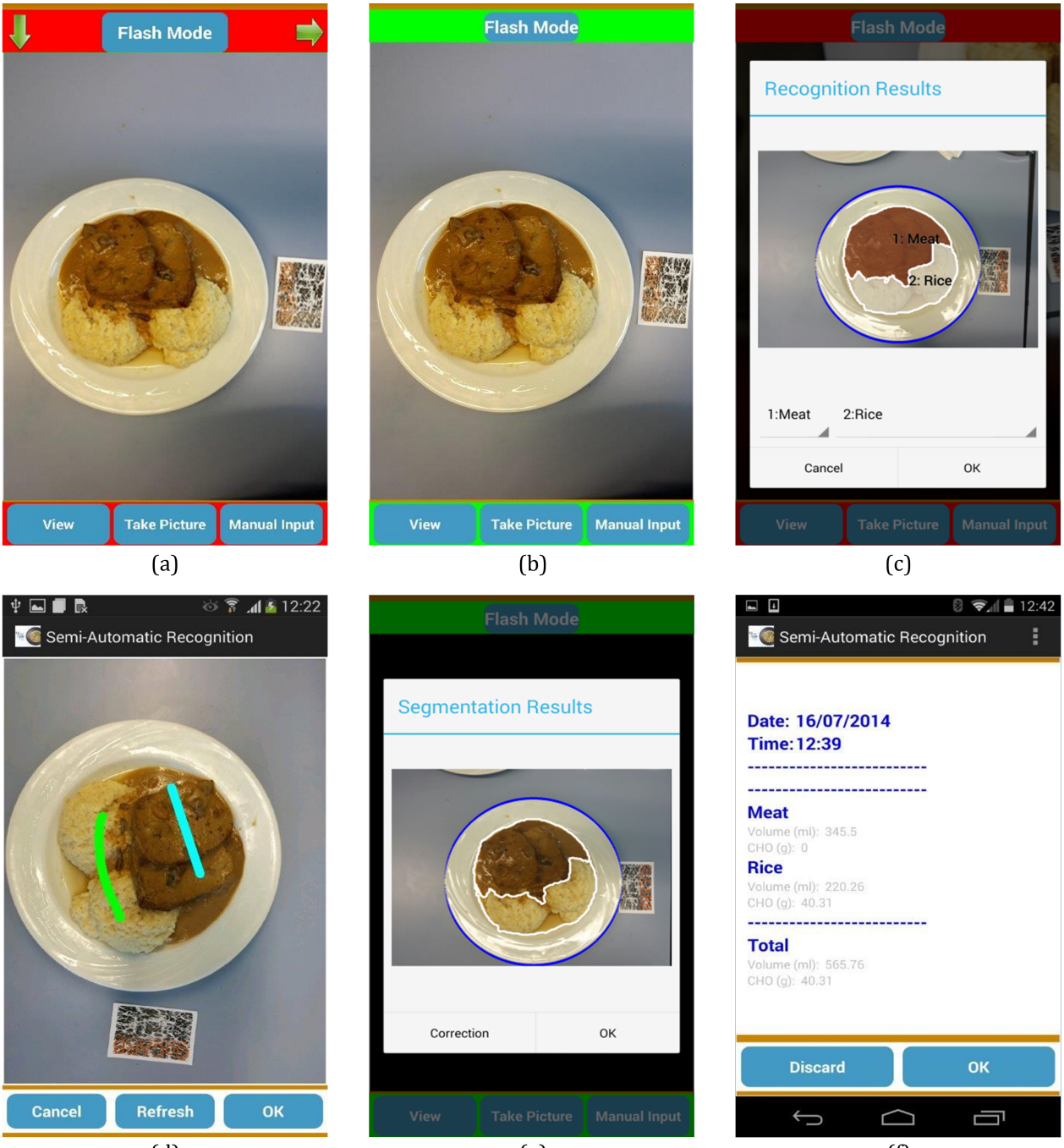


Figure 3. Screenshots of the GoCARB app: (a) the red frame indicates wrong angle so image acquisition is disabled, (b) when the frame turns green, the user can take an image, (c) the result of the automatic segmentation, (d) user-given seeds required for the semiautomatic segmentation, (e) the result of the recognition with the option of manual correction, and (f) the results screen displayed to the user.



Study Setup

Nineteen adult individuals receiving regular care at the Division of Endocrinology, Diabetes and Clinical Nutrition of the Bern University Hospital, "Inselspital" (Bern, Switzerland) were asked to participate in the GoCARB study, which lasted a total of 10 days during July and August 2014. The GoCARB system was preinstalled on two types of mobile phones, Samsung Galaxy S4 and Nexus 5, both running the Android operation system (version Jelly Bean 4.2). Every day of the study, two standard dishes were ordered from the hospital restaurant, each in three sizes (small, normal, and large), resulting in six different dishes per day and 60 dishes in total. Each dish contained three food items that corresponded to common sources of protein (eg, meat, fish), carbohydrates (eg, pasta, rice), and vegetables/salads (eg, lettuce, carrots). Figure 4 presents some examples of these dishes. The food items were weighed with a household scale and the true carbohydrate content in grams was defined using the USDA National Nutrient Database for Standard Reference. Each participant joined the study for one session during which she/he was asked to estimate the carbohydrate content of each of the six meals on her/his own and then by using the GoCARB system. The 19 participants were randomly distributed over the 10 days of the study so some dishes were shown to multiple

participants. Therefore, a total of 114 (19×6) estimations were made even though there were 60 unique dishes. Before using the app, every participant received short training and a detailed written user manual. At the end of each session, a questionnaire was completed to assess the user's experience with GoCARB. The questionnaire was a combination of closed and open questions and was used to gather information about the satisfaction and perceived usefulness of the mobile phone app. Because this study did not collect nor analyze patient-specific clinical data, the study was exempted from formal ethical approval.

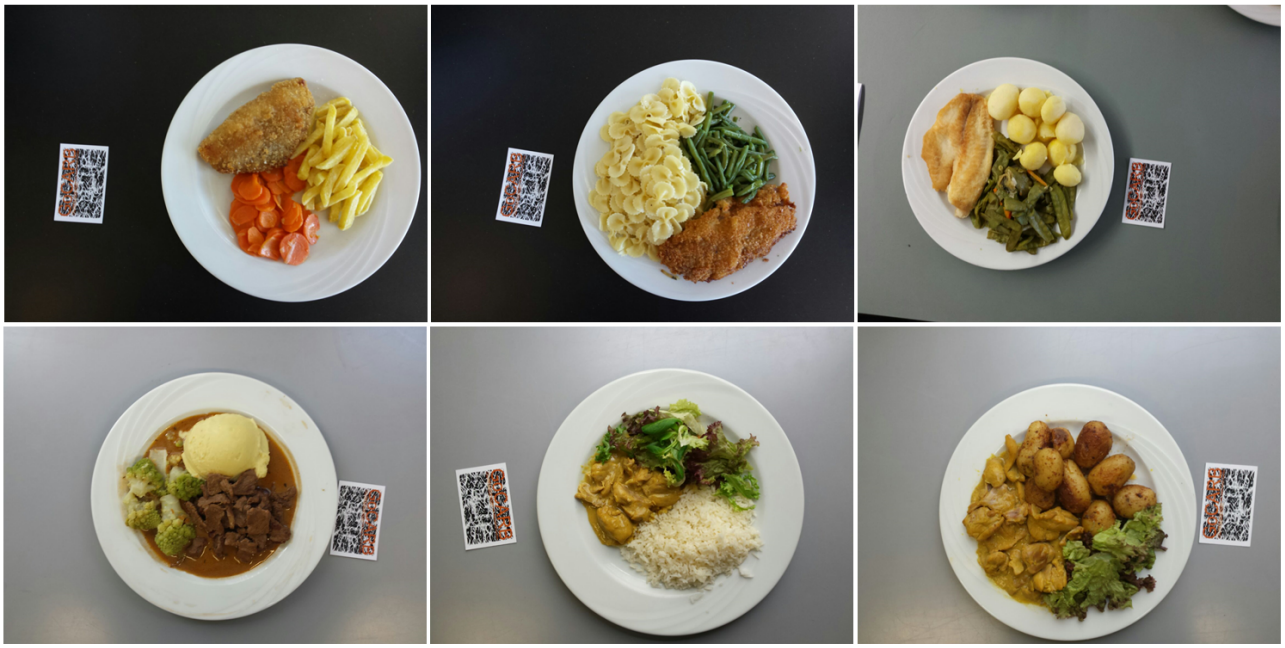
Participants

The mean age of the participants was 40.5 (SD 11.5) years and all but one were mobile phone users. Among the mobile phone owners, 12 were familiar with iOS, five with Android, and one with Blackberry OS. Of these, 16% (3/19) used a nutrition-related app in their everyday life, whereas only 11% (2/19) used a health-related app.

Statistical Analysis

For statistical comparison, we used the mean absolute error of the participants' estimates with and without using GoCARB. The data were nonnormally distributed; therefore, an independent Mann-Whitney *U*test was applied for significance testing. Descriptive statistical analysis was performed in open-source software R [26] and SPSS version 23 (IBM Corp, Armonk, NY, USA). The questionnaire investigating usability and user satisfaction included 24 Likert-scale questions, three polar questions, and three open questions. Likert-scale questions ranged from "do not agree at all," "do not agree," "neutral," "agree," to "agree strongly" for general GoCARB usability and from "very bad," "bad," "neutral," "good," to "very good" for the performance of automatic and interactive GoCARB features.

Figure 4. Examples of the dishes used for the study.



Results

Carbohydrate Estimation: GoCARB Versus Individuals With Type 1 Diabetes

As presented in Table 1, the mean absolute error of the participants with type 1 diabetes was 27.89 (SD 38.20) grams of carbohydrate, whereas the corresponding values for the GoCARB system was 12.28 (SD 9.56) grams of carbohydrate, less than the initial target of 20 grams. This improvement in the estimation error was statistically significant (*P*=.001). The corresponding mean relative error in carbohydrate estimation was 54.8% (SD 72.3%) for the participants and 26.2% (SD 18.7%) for GoCARB.

Table 1. Performance of participants (N=19) in carbohydrate estimation with and without GoCARB.

All participants	Absolute error (grams), mean (SD)	Absolute percentage error (%), mean (SD)	Absolute errors <20 grams, n (%)
Without GoCARB	27.89 (38.20)	54.8 (72.3)	67/114 (58.8)
With GoCARB	12.28 (9.56)	26.2 (18.7)	92/114 (80.7)

Distribution of Estimation Errors

Figure 5 provides the distribution of errors for carbohydrates as counted by the individuals with type 1 diabetes and as estimated by using the GoCARB system. In the case of self-assessment, the error distribution was broad with outliers up to 200 grams of carbohydrate. In the case of GoCARB, the errors were symmetric and concentrated around zero. In general, the individuals with type 1 diabetes more frequently (60.5%, 69/114) underestimated the carbohydrates of a meal; however,

the error was higher in the case of the overestimation. The GoCARB system exhibited an evenly balanced distribution of under- and overestimation, 50.9% (58/114) and 49.1% (56/114), respectively, whereas the errors were almost of the same magnitude for both under- and overestimations. Furthermore, in 58.8% (67/114) of the cases of carbohydrates counted by the individuals with type 1 diabetes, the error was in the range of –20 grams to +20 grams, whereas with GoCARB the estimations were in the required range in 80.7% (92/114) of the cases. The analysis was followed by looking for outliers in carbohydrate estimations as counted by the individuals with type 1 diabetes. Examining the data (see Figure 6), one participant was identified who consistently overestimated all her/his meals. The mean absolute estimation error for this participant was 158.19 (SD 26.09) grams of carbohydrate. As seen in Table 2, even by excluding the one participant with extreme carbohydrate values, the use of GoCARB resulted in significantly better performances (P=.01).

Table 2. Performance of participants in carbohydrate estimation with and without GoCARB after excluding one participant with extreme errors (n=18).

Excluding extremely bad estimator	Absolute error (grams), mean (SD)	Absolute percentage error (%), mean (SD)	Absolute errors <20 grams, n (%)
Without GoCARB	17.81 (14.94)	34.3 (24.3)	67/108 (62.0)
With GoCARB	12.75 (9.84)	26.9 (18.9)	86/108 (79.6)

Figure 5. Distribution of the absolute errors in carbohydrate estimation with and without GoCARB.

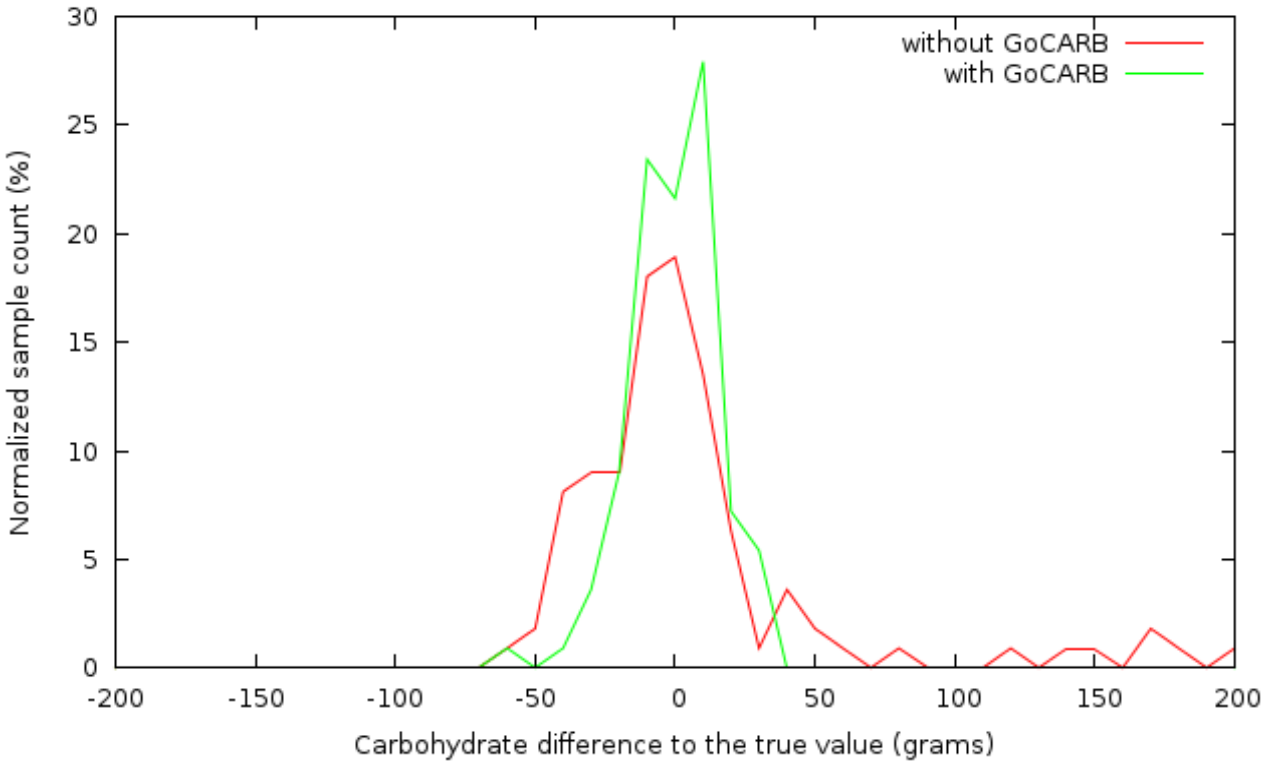
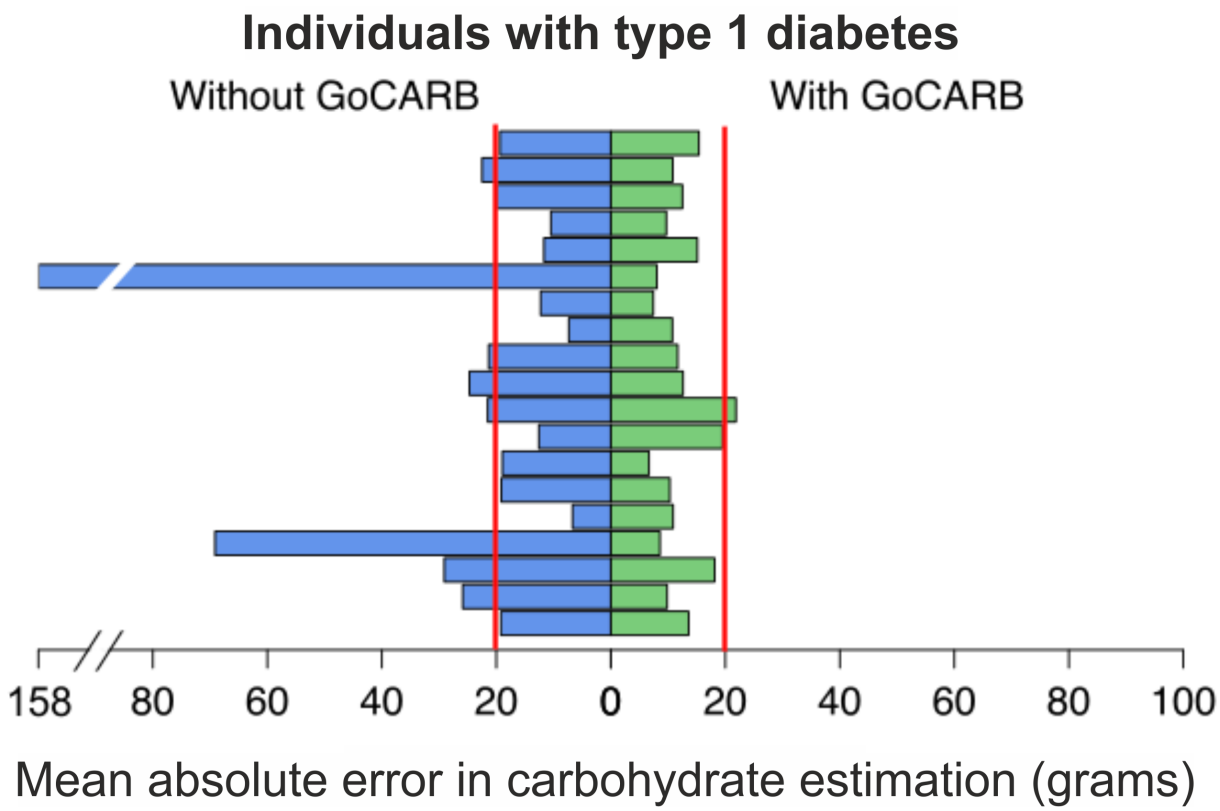


Figure 6. Distribution of the mean absolute error per participant with and without GoCARB. The vertical lines represent the target range of errors.



Influence of Meal Size on Accuracy

For each meal, three different meal sizes were presented to the user. With larger meals, individual's counting errors also increased: mean 18.47 (SD 28.86) grams of carbohydrate for small meals, mean 26.39 (SD 38.20) grams of carbohydrate for medium meals, and mean 38.82 (SD 47.03) grams of carbohydrate for large meals. When the participants used the GoCARB system to estimate the carbohydrate content of the meals, the error exhibited less variation across the different meal sizes: mean 10.34 (SD 6.16) grams of carbohydrate for small, mean 10.12 (SD 8.02) grams of carbohydrate for medium, and mean 16.38 (SD 12.30) grams of carbohydrate for large-sized meals.

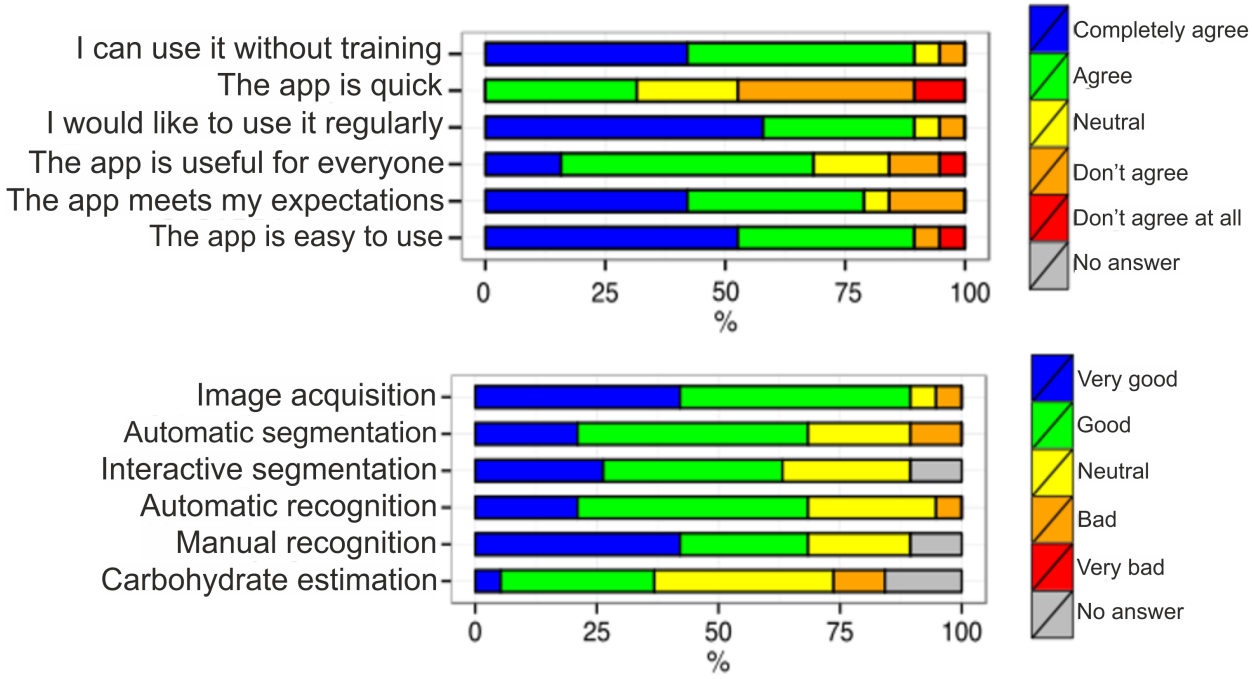
Usage of Interactive Features

In the GoCARB prototype, the user is able to interact with the system when she/he is not satisfied with the fully automatic segmentation and/or recognition results. Thus, the system can be corrected, while the volume estimation is completed. During the study, the automatic segmentation was successful in 75.4% (86/114) of the cases. For recognition, the system achieved correct recognition of all three distinct food items on a dish in 59.6% (68/114) of cases. For these cases, manual correction was not needed. In 36.0% (41/114) of cases, the system was able to recognize two of three food items correctly, whereas only one food item was recognized in 4.4% (5/114) of cases. For these cases, the user had to manually correct the system. In all cases, GoCARB recognized correctly at least one of the food items. Out of a total of 342 food items in 114 meals, 85.1% (291/342) were correctly automatically recognized.

User Experience

To evaluate the user perception and usability aspects of GoCARB, a questionnaire was given to participants. The majority (90%, 17/19) agreed or agreed completely that GoCARB was easy to use (Figure 7). Furthermore, 90% (17/19) of participants would have liked to use GoCARB on a regular basis and 68% (13/19) thought GoCARB was useful for all individuals with type 1 diabetes. The processing speed of the GoCARB system was considered too slow by 47% (9/19). It is noteworthy that the client-server architecture of the system made its speed highly dependent on 3/4G or Wi-Fi signal strength. The GoCARB workflow was mostly rated favorable, with moderate to high approval of automated and interactive steps. The feedback from the open questions was similar. The participants were asked to list negative and positive aspects of the app and possibly to provide suggestions for improving the system. Most of the negative feedback was related to processing delays and the dependence on the Internet connection. As for positive aspects, the participants indicated the ease of use and the usefulness of such an app. Many participants proposed that the system should be extended to other food types, such as desserts.

Figure 7. Graphical representation of the participants' answers to the questionnaire.



Discussion

Principal Findings

The results of the study indicate that GoCARB is, on average, more accurate at estimating carbohydrate content than the individuals with type 1 diabetes who participated. If all participants are included, the mean absolute estimation error while using GoCARB (12.28 grams of carbohydrate) was reduced by more than 50% than without using GoCARB (27.89 grams of carbohydrate). If it is considered that an absolute error of 20 grams or more gives a significant increase in risk of hyper- or hypoglycemia, GoCARB fulfilled this prerequisite for 80.7% (92/114) of the cases. Moreover, the relatively low standard deviation (SD 9.56 grams) of GoCARB's errors, along with their symmetric distribution around zero, demonstrates the relative stability and consistency of the system. In 75.4% (86/114) of the meals, GoCARB's automatic segmentation was successful and 85.1% (291/342) of individual food items were successfully recognized. With regard to user satisfaction, the majority of participants found the system easy to use and expressed strong interest in using it on a daily basis. Among these individuals with type 1 diabetes, there was a strong consensus that an app such as GoCARB could benefit everybody with diabetes. The only concern of the participants was related to the speed of the app (approximately 1 min including user interactions with the system), which was mainly due to delays in the transmission of data over an unstable wireless network.

The errors in carbohydrate estimation by the participants exhibited a wide spread, but only one participant was responsible for all the estimates beyond the 95th percentile. Despite different approaches to outlier exclusion, the overall difference between individuals with type 1 diabetes with and without GoCARB remains significant. Whether it is reasonable to include outliers in our statistical analysis greatly depends on the point of view. Although statistical analysis permits outlier exclusion, clinical reasoning and experience supports their inclusion, especially because everyday clinical life shows that extremes are possible and might even be the accidental starting point of glycemic decompensation, in extremis leading to hospitalization or death. Undoubtedly, these diabetics with weak estimation skills would benefit the most. Nevertheless, even after excluding this participant, the difference in accuracy between the conventional methods and GoCARB remains statistically significant (P=.01).

Comparison With Previous Work

There have been several systems proposed using single/multiple food images or videos to recognize food items and calculate the corresponding volumes [14-21]. However, to the best of our knowledge, there has been no comparable study evaluating an automatic dietary assessment system together with end users. From our experience with the GoCARB system, we believe that a thorough scientific evaluation and an appropriate framework for mobile medical app development and evaluation are of the utmost importance. Only a few commercially available medical apps undergo in-depth scientific evaluation and there is virtually no support provided to health care professionals or patients to help them to select the most beneficial app. This concern is increasingly being addressed and steps have been taken to develop guidelines and standards [27,28].

Limitations

This is a preclinical study and the sample number was chosen for reasons of convenience. Within the study sample, there was a wide spread of estimation error and estimation error was significantly higher than that reported in the literature. Nevertheless, we believe that this is concordant with many patient groups found in real life, but only studies with increased sample size may bring certainty. Moreover, we do not know if or when participants last had a nutritionist-guided training

session on carbohydrate counting. Because there is general agreement that regular teaching and training significantly improve estimation capability [29], such information is of interest, especially to identify individuals with type 1 diabetes who perform poorly despite adequate training.

Furthermore, purely automatic computer vision-based systems have certain limitations. For complex meal types with multiple ingredients mixed arbitrarily or meals covered by sauce, additional information is needed by the user on the food type (eg, lasagne). The rest of the modules are independent of food type. Macronutrients, such as fat, are nearly impossible to quantify by computer vision only, but can still affect a person's postprandial glycemia. Moreover, the type of carbohydrate also affects the postprandial response, apart from its amount. Intrinsic variables that influence the effect of carbohydrate-containing meals on blood glucose response include the specific type of food ingested, type of starch (amylose vs amylopectin), style of preparation (cooking method and time, amount of heat or moisture used), ripeness, and degree of processing. To some extent, this is reflected in the glycemic index of foods and is an established method to compare the physiological postprandial glucose responses to different types of carbohydrate-containing foods [30]. Essential components of a system such as GoCARB are the multimedia and nutritional databases used for recognizing the food and calculating its nutritional profile. Generating and expanding a food image dataset to provide a broad variety for automatic food recognition is a challenging and often costly task. In addition, although technology has improved the accessibility of nutritional databases, the availability of up-to-date food composition data for many food items is still limited. Although new products continually appear in the food supply, gaps will always exist between what databases contain and what individuals consume.

Future Research

We plan to optimize the existing prototype and expand its functionalities. The code of the system will be optimized and components will be moved from the server to the mobile phone side to increase its speed and efficiency. The food types considered will be expanded to cover a wider spectrum of cultures and eating habits. Integration of a barcode reader would be easily feasible technically and would allow us to cover packaged foods. Other input methods are promising; these might include voice input or dropdown lists specifying details of a meal that cannot be assessed by a picture (eg, fried vs boiled, cooked vs uncooked, olive oil vs sunflower oil).

Although GoCARB aims to support individuals with type 1 diabetes, the system might be extended so that it not only displays carbohydrates, but also other micro- and macronutrients. The system could then be used to facilitate diet management. Such a system could be useful for a large heterogeneous group of medical conditions related to food, food intake, and digestion or weight management. It could also be aimed toward a more general and healthy population in an attempt to encourage people to make more deliberate food choices leading to a healthier lifestyle. Such a tool would not only be of interest to individuals, but also to the nutrition research community because conventional methods of diet assessment, such as the 24-hour food recall method, food intake questionnaires, or a paperback food diary, are difficult to apply, are time consuming, and are known to be consistently inaccurate.

Conclusion

In this study, we have presented the evaluation of a novel mobile phone-based system to estimate the carbohydrate content of a meal on a plate using computer vision. The GoCARB system proves to be a reliable support tool for carbohydrate estimation of meals on a plate and provides more accurate carbohydrate estimates than those of our cohort of participants with type 1 diabetes using conventional methods. Among the participants, there was a strong consensus that an app such as the GoCARB system could benefit everybody with diabetes. We believe that computer vision has the potential to facilitate and ameliorate the cumbersome and error-prone task of carbohydrate estimation. The output of GoCARB could be used as input to a bolus calculator that will also consider the personal characteristics of the user (eg, insulin-to-carbohydrate ratio) and suggest an insulin dose. After further development, GoCARB could ultimately make a distinct contribution to a fully automated artificial pancreas.

Acknowledgments

We would like to thank the volunteers for participating in the GoCARB study and Swisscom for generously providing the mobile phones used. The work was funded in part by the Bern University Hospital "Inselspital" and the European Union Seventh Framework Programme (FP7-PEOPLE-2011-IAPP) under grant agreement no. 286408.

Conflicts of Interest

None declared.

References

1. International Diabetes Federation. 2015. IDF Diabetes Atlas, 7th Edition URL: http://www.diabetesatlas.org/ [accessed 2016-04-26] [WebCite Cache ID 6h3PfB5wn]
2. Guariguata L, Nolan T, Beagley J, Linnenkamp U, Jacqmain O. IDF Diabetes Atlas, Sixth Edition. Brussels: International Diabetes Foundation; 2013. URL: http://www.idf.org/sites/default/files/EN_6E_Atlas_Full_0.pdf [accessed 2016-04-26] [WebCite Cache ID 6h3Rxi8Nt]

3. Forouhi NG, Wareham NJ. Epidemiology of diabetes. Medicine (Abingdon) 2014 Dec;42(12):698-702 [FREE Full text] [doi: 10.1016/j.mpmed.2014.09.007] [Medline: 25568613]
4. Smart CE, Ross K, Edge JA, Collins CE, Colyvas K, King BR. Children and adolescents on intensive insulin therapy maintain postprandial glycaemic control without precise carbohydrate counting. Diabet Med 2009 Mar;26(3):279-285. [doi: 10.1111/j.1464-5491.2009.02669.x] [Medline: 19317823]
5. Smart CE, King BR, McElduff P, Collins CE. In children using intensive insulin therapy, a 20-g variation in carbohydrate amount significantly impacts on postprandial glycaemia. Diabet Med 2012 Jul;29(7):e21-e24. [doi: 10.1111/j.1464-5491.2012.03595.x] [Medline: 22268422]
6. Brazeau AS, Mircescu H, Desjardins K, Leroux C, Strychar I, Ekoé JM, et al. Carbohydrate counting accuracy and blood glucose variability in adults with type 1 diabetes. Diabetes Res Clin Pract 2013 Jan;99(1):19-23. [doi: 10.1016/j.diabres.2012.10.024] [Medline: 23146371]
7. Bishop FK, Maahs DM, Spiegel G, Owen D, Klingensmith GJ, Bortsov A, et al. The Carbohydrate Counting in Adolescents with Type 1 diabetes (CCAT) study. Diabetes Spectrum 2009 Jan 01;22(1):56-62. [doi: 10.2337/diaspect.22.1.56]
8. Smart CE, Ross K, Edge JA, King BR, McElduff P, Collins CE. Can children with Type 1 diabetes and their caregivers estimate the carbohydrate content of meals and snacks? Diabet Med 2010 Mar;27(3):348-353. [doi: 10.1111/j.1464-5491.2010.02945.x] [Medline: 20536499]
9. Graff MR, Gross TM, Juth SE, Charlson J. How well are individuals on intensive insulin therapy counting carbohydrates? Diabetes Res Clin Pr 2000 Sep;50:238-239. [doi: 10.1016/S0168-8227(00)82267-6]
10. American Diabetes Association. Executive summary: standards of medical care in diabetes--2014. Diabetes Care 2014 Jan;37 Suppl 1:S5-S13. [doi: 10.2337/dc14-S005] [Medline: 24357214]
11. Rusin M, Arsand E, Hartvigsen G. Functionalities and input methods for recording food intake: a systematic review. Int J Med Inform 2013 Aug;82(8):653-664. [doi: 10.1016/j.ijmedinf.2013.01.007] [Medline: 23415822]
12. Martin CK, Nicklas T, Gunturk B, Correa JB, Allen HR, Champagne C. Measuring food intake with digital photography. J Hum Nutr Diet 2014 Jan;27 Suppl 1:72-81 [FREE Full text] [doi: 10.1111/jhn.12014] [Medline: 23848588]
13. Domhardt M, Tiefengrabner M, Dinic R, Fötschl U, Oostingh GJ, Stütz T, et al. Training of carbohydrate estimation for people with diabetes using mobile augmented reality. J Diabetes Sci Technol 2015 May;9(3):516-524. [doi: 10.1177/1932296815578880] [Medline: 25883165]
14. He H, Kong F, Tan J. DietCam: multi-view food recognition using a multi-kernel SVM. IEEE J Biomed Health Inform 2015 Apr 2:e1. [doi: 10.1109/JBHI.2015.2419251] [Medline: 25850095]
15. Pouladzadeh P, Shirmohammadi S, Arici T. Intelligent SVM based food intake measurement system. 2013 Presented at: IEEE International Conference on Computational Intelligence and Virtual Environments for Measurement Systems and Applications (CIVEMSA); Jul 15-17, 2013; Milan, Italy. [doi: 10.1109/CIVEMSA.2013.6617401]
16. Anthimopoulos M, Dehais J, Shevchik S, Ransford BH, Duke D, Diem P, et al. Computer vision-based carbohydrate estimation for type 1 patients with diabetes using smartphones. J Diabetes Sci Technol 2015 May;9(3):507-515. [doi: 10.1177/1932296815580159] [Medline: 25883163]
17. Shroff G, Smailagic A, Siewiorek D. Wearable context-aware food recognition for calorie monitoring. In: Proceedings of the 2008 12th IEEE International Symposium on Wearable Computers. 2008 Presented at: ISWC '08 12th IEEE International Symposium on Wearable Computers; Sep 28-Oct 1, 2008; Pittsburgh, PA p. 119-120. [doi: 10.1109/ISWC.2008.4911602]
18. Chen M, Dhingra K, Wu W, Yang L, Sukthankar R, Yang J. PFID: Pittsburgh fast-food image dataset. In: Proceedings of the 16th IEEE International Conference on Image Processing. 2009 Presented at: ICIP'09 16th IEEE International Conference on Image Processing; Nov 7-10, 2009; Cairo, Egypt p. 289-292. [doi: 10.1109/ICIP.2009.5413511]
19. Wu W, Yang J. Fast food recognition from videos of eating for calorie estimation. In: Proceedings of the 2009 IEEE International Conference on Multimedia and Expo. 2009 Presented at: ICME'09 2009 IEEE International Conference on Multimedia and Expo; Jun 28-Jul 3, 2009; New York p. 1210-1213. [doi: 10.1109/ICME.2009.5202718]
20. Zhu F, Bosch M, Woo I, Kim S, Boushey CJ, Ebert DS, et al. The use of mobile devices in aiding dietary assessment and evaluation. IEEE J Sel Top Signal Process 2010 Aug;4(4):756-766 [FREE Full text] [doi: 10.1109/JSTSP.2010.2051471] [Medline: 20862266]
21. Puri M, Zhu Z, Yu Q, Divakaran A, Sawhney H. Recognition and volume estimation of food intake using a mobile device. In: Applications of Computer Vision (WACV), 2009 Workshop on. 2009 Presented at: Workshop on Applications of Computer Vision (WACV); Dec 7-8, 2009; Snowbird, UT p. 1-8. [doi: 10.1109/WACV.2009.5403087]
22. Dehais J, Anthimopoulos M, Mougiakakou S. Dish detection and segmentation for dietary assessment on smartphones. In: New Trends in Image Analysis and Processing--ICIAP 2015 Workshops. 2015 Sep Presented at: ICIAP 2015 International Workshops; Sep 7-8, 2015; Genoa, Italy p. 433-440. [doi: 10.1007/978-3-319-23222-5_53]
23. Anthimopoulos M, Dehais J, Diem P, Mougiakakou S. Segmentation and recognition of multi-food meal images for carbohydrate counting. In: Bioinformatics and Bioengineering (BIBE), 2013 IEEE 13th International Conference on. 2013 Presented at: IEEE 13th International Conference on Bioinformatics and Bioengineering (BIBE); Nov 10-13, 2013; Chania, Greece. [doi: 10.1109/BIBE.2013.6701608]
24. Dehais J, Shevchik S, Diem P, Mougiakakou S. Food volume computation for self dietary assessment applications. In: Bioinformatics and Bioengineering (BIBE), 2013 IEEE 13th International Conference on. 2013 Presented at: IEEE 13th

International Conference on Bioinformatics and Bioengineering (BIBE); Nov 10-13, 2013; Chania, Greece p. 1. [doi: 10.1109/BIBE.2013.6701615]
25. USDA. 2015. USDA National Nutrient Database for Standard Reference URL: https://ndb.nal.usda.gov/ [accessed 2016-04-26] [WebCite Cache ID 6h3UW61MR]
26. R Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2014. URL: http://www.R-project.org/ [accessed 2016-04-26] [WebCite Cache ID 6h3VkiJ6p]
27. Powell AC, Landman AB, Bates DW. In search of a few good apps. JAMA 2014 May 14;311(18):1851-1852. [doi: 10.1001/jama.2014.2564] [Medline: 24664278]
28. Mobile Medical Applications: Guidance for Industry and Food and Drug Administration Staff. Rockville, MD: US Department of Health and Human Services Food and Drug Administration; 2015 Feb 09. URL: http://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/UCM263366.pdf [accessed 2016-04-26] [WebCite Cache ID 6h3VzfTIM]
29. Spiegel G, Bortsov A, Bishop FK, Owen D, Klingensmith GJ, Mayer-Davis EJ, et al. Randomized nutrition education intervention to improve carbohydrate counting in adolescents with type 1 diabetes study: is more intensive education needed? J Acad Nutr Diet 2012 Nov;112(11):1736-1746 [FREE Full text] [doi: 10.1016/j.jand.2012.06.001] [Medline: 22975086]
30. American Diabetes Association, Bantle JP, Wylie-Rosett J, Albright AL, Apovian CM, Clark NG, et al. Nutrition recommendations and interventions for diabetes: a position statement of the American Diabetes Association. Diabetes Care 2008 Jan;31 Suppl 1:S61-S78. [doi: 10.2337/dc08-S061] [Medline: 18165339]

Abbreviations

USDA: US Department of Agriculture

Edited by G Eysenbach; submitted 26.01.16; peer-reviewed by S Bromuri, M Tiefengrabner, M Allman-Farinelli; comments to author 14.02.16; revised version received 09.03.16; accepted 21.03.16; published 11.05.16.

Please cite as:
Rhyner D, Loher H, Dehais J, Anthimopoulos M, Shevchik S, Botwey RH, Duke D, Stettler C, Diem P, Mougiakakou S
Carbohydrate Estimation by a Mobile Phone-Based System Versus Self-Estimations of Individuals With Type 1 Diabetes Mellitus: A Comparative Study
J Med Internet Res 2016;18(5):e101
URL: http://www.jmir.org/2016/5/e101/
doi:10.2196/jmir.5567
PMID:27170498

Original Paper

The Effectiveness of Self-Management Mobile Phone and Tablet Apps in Long-term Condition Management: A Systematic Review

Lisa Whitehead[1*], BSc (Hons), MA, PhD; Philippa Seaton[2*], BA, MA (Hons), PhD

[1]School of Nursing and Midwifery, Edith Cowan University, Joondalup, Australia

[2]Centre of Postgraduate Nursing Studies, University of Otago, Christchurch, New Zealand

[*]all authors contributed equally

Corresponding Author:
Lisa Whitehead, BSc (Hons), MA, PhD
School of Nursing and Midwifery
Edith Cowan University
Building 21
270 Joondalup Drive
Joondalup, 6027
Australia
Phone: 61 (8)63042641
Fax: 61 (8)63042323
Email: l.whitehead@ecu.edu.au

Abstract

Background: Long-term conditions and their concomitant management place considerable pressure on patients, communities, and health care systems worldwide. International clinical guidelines on the majority of long-term conditions recommend the inclusion of self-management programs in routine management. Self-management programs have been associated with improved health outcomes; however, the successful and sustainable transfer of research programs into clinical practice has been inconsistent. Recent developments in mobile technology, such as mobile phone and tablet computer apps, could help in developing a platform for the delivery of self-management interventions that are adaptable, of low cost, and easily accessible.

Objective: We conducted a systematic review to assess the effectiveness of mobile phone and tablet apps in self-management of key symptoms of long-term conditions.

Methods: We searched PubMed, Embase, EBSCO databases, the Cochrane Library, and The Joanna Briggs Institute Library for randomized controlled trials that assessed the effectiveness of mobile phone and tablet apps in self-management of diabetes mellitus, cardiovascular disease, and chronic lung diseases from 2005–2016. We searched registers of current and ongoing trials, as well as the gray literature. We then checked the reference lists of all primary studies and review papers for additional references. The last search was run in February 2016.

Results: Of the 9 papers we reviewed, 6 of the interventions demonstrated a statistically significant improvement in the primary measure of clinical outcome. Where the intervention comprised an app only, 3 studies demonstrated a statistically significant improvement. Interventions to address diabetes mellitus (5/9) were the most common, followed by chronic lung disease (3/9) and cardiovascular disease (1/9). A total of 3 studies included multiple intervention groups using permutations of an intervention involving an app. The duration of the intervention ranged from 6 weeks to 1 year, and final follow-up data ranged from 3 months to 1 year. Sample size ranged from 48 to 288 participants.

Conclusions: The evidence indicates the potential of apps in improving symptom management through self-management interventions. The use of apps in mHealth has the potential to improve health outcomes among those living with chronic diseases through enhanced symptom control. Further innovation, optimization, and rigorous research around the potential of apps in mHealth technology will move the field toward the reality of improved health care delivery and outcomes.

(J Med Internet Res 2016;18(5):e97) doi:10.2196/jmir.4883

KEYWORDS

mobile phone; apps; telemedicine; mHealth; self-management; chronic disease; diabetes mellitus; cardiovascular diseases; lung diseases; systematic review

Introduction

The number of people living with 1 or more chronic disease continues to increase worldwide [1]. Improvement in living conditions and treatment has increased the life expectancy of people with chronic conditions; however, without effective management, quality of life may be poor. Empowering and engaging people with chronic diseases to manage their own health is vital. Several barriers have been identified in the ability of the individual and health care providers to work together to promote self-management [2]. Having access to timely information, assessment, and treatment are all vital in the management of long-term conditions [3]. mHealth interventions offer the potential to overcome many of the traditional barriers by offering convenience and care in a natural environment and minimizing the barriers of distance, time, and cost. For the clinician, mHealth interventions offer the ability to evaluate a prescribed course of action, monitor adverse events, and identify areas for improvement [4,5].

For the past decade, mHealth has been constantly expanding as a subdivision of eHealth. Mobile apps for health have the potential to target heterogeneous populations, but with the ability to also address specific needs and complement highly developed health care technologies. The market is evolving rapidly, generating myriad opportunities for the development of new mobile technologies [6].

Mobile phones (ie, mobile phones with advanced computing and Internet access) and tablet computers (ie, general purpose computers contained in a single panel and usually operated through a touch screen) have become the most popular and widespread types of mobile device [7]. Close to 55% of British adults claim to own a mobile phone [8] and over a third own a tablet [9]. In the United States, a report by the Pew Research Center found that 64% of all adults now own a mobile phone [10] and 34% of American adults own a tablet computer [11]. Worldwide, just under 17% of the 6 billion mobile subscriptions are mobile phone subscriptions [12]. As retail prices decline, ownership of these devices is likely to continue to increase [8] in high-income and low- and middle-income countries.

Mobile apps are increasingly used in managing various tasks in daily life. More than 900,000 apps are available in the Apple App Store (iOS operating system; Apple Inc) and more than 700,000 apps in the Google Play Store (Android operating system; Google). Over 100,000 of these are health-related apps.

Sophisticated computing features mean that both mobile phones and tablet computers can support self-management functions and deliver them at a population level. Self-management interventions could be offered within software extensions that users add to their devices, popularized under the term apps [13].

Methods

Overview

We undertook a systematic review of apps used to facilitate self-management of long-term conditions with an outcome focus on the key disease markers and symptoms. The long-term conditions were diabetes mellitus, cardiovascular diseases, and chronic lung diseases. Cardiovascular diseases include hypertension, coronary artery disease, and congestive heart failure. Chronic lung diseases include asthma and chronic obstructive pulmonary disease. We chose these conditions on account of their high global burden [14].

Our definition of self-management apps was software programs designed for mobile phones and tablets that aim to promote or support self-management skills to manage the key disease markers and symptoms. Apps are optional add-ons to the device that interact with users through a set of interfaces (eg, a visual user interface). Health apps can be characterized as a medium with broad capabilities to communicate information, provide interactive experiences, and collect information from patients. They provide a platform for the delivery of self- management interventions that are highly adaptable, of low cost to the health system, and easily accessible.

Using Boolean phrases, we searched PubMed, Embase and EBSCO databases for studies that assessed the effectiveness of apps in the management of diabetes, cardiovascular diseases, and chronic lung diseases. We searched PubMed using Medical Subject Headings and advanced search builder features. Emtree terms using the explosion function to extend the search were used to build a multiterm query along with advanced searches in Embase. We included CINAHL, PsycINFO, and PsycARTICLES in the EBSCO database search. We handsearched the JMIR journals and *Telemedicine Journal and e-Health* as the key publications for this area of research. We examined the reference lists of all papers included in the review and removed duplicates. The databases were searched between 2005 and 2016, since technologies prior to 2005 are unlikely to be representative of contemporary technologies that support health apps [15-17]. In addition, the concept of self-management was not widely adopted prior to 2005. We completed the final search in February 2016 (Textbox 1). The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement guided the reporting of the review [18].

Textbox 1. PubMed search strategy (terms).

1. Pulmonary disease, chronic obstructive
2. asthma
3. Acute coronary syndrome
4. Blood glucose
5. Blood pressure
6. Coronary Disease
7. Cardiovascular diseases
8. Diabetes mellitus
9. Forced expiratory volume (FEV)
10. Hemoglobin A, glycosylated
11. Hypertension
12. Computers, Handheld
13. Peak expiratory flow rate
14. Cell Phones
15. MP3-Player
16. Telemedicine
17. Or/1-16
18. Limit 17 to yr=2005-Current

Inclusion and Exclusion Criteria

We included original research published in peer reviewed journals that evaluated self-management apps for their effect on disease-specific clinical outcomes. The focus on disease-specific clinical measures such as glycated hemoglobin (HbA_{1c}) or blood pressure was chosen because improved clinical outcomes are the ultimate goal (in terms of quality-adjusted life years, disease burden, and health care costs) of self-management programs. Included studies were randomized controlled trials (RCTs) of self-management interventions for patients with a clinician-diagnosed long-term condition delivered via mobile phone apps compared with either self-management interventions delivered via traditional methods (eg, paper-based diaries) or usual care.

We excluded papers if (1) they reported on primary prevention among healthy or at-risk groups, (2) the focus lay outside of the self-management domain (see Textbox 2 for self-management-related activities [3,19]), (3) the sample did not include people living with diabetes, cardiovascular diseases, or chronic lung diseases or where the results from the subsample of the populations of interest were not distinctly reported, or (4) the intervention targeted health care professionals; required modification of hardware; relied solely on messaging (short message service or multimedia message service); did not offer a mode of interaction (this could be automated and based on logarithms), acting only as a transmitter of data (eg, from patient to clinician), because this would be more reflective of telemonitoring; or used devices that did not offer portability comparable with mobile phones and tablets (eg, desktops, laptops, notebooks, and netbooks)—although these are portable, they are not accessible at all times regardless of location. In addition, we excluded review papers, editorials, commentaries, dissertations, poster presentations, abstracts only, proposals for future studies, study protocols, and descriptive papers describing apps but not testing them in a sample population. Publication language was restricted to English only.

Textbox 2. Patients' self-management characteristics (adapted from Lahdensuo [19] and Battersby et al [3]).

- Accept the condition as a long-term disease amenable to intervention
- Have knowledge about the disease and its treatment
- Actively participate in the control and management of the disease
- Identify factors that make the condition worse
- Be able to describe strategies for avoidance or reduction of exacerbating factors
- Recognize the signs and symptoms of deterioration in health
- Follow a prescribed, written treatment plan
- Use correct technique for taking drugs
- Take appropriate action to prevent and treat symptoms in different situations
- Use medical resources appropriately for routine and acute care
- Monitor symptoms and objective measures of disease control
- Identify barriers to adherence to the treatment plan
- Address specific problems that have an impact on the individual's condition

Data Extraction and Analysis

We initially screened publications for potential inclusion based on simultaneous review of title and abstract by 2 reviewers. Any discrepancies were resolved by consensus between the reviewers with reference to the full paper. Information extracted from each paper using a structured form included objectives, types of intervention, setting, sample characteristics, outcomes measured, and results reported. We assessed risk of bias for all included studies using the Cochrane Collaboration's tool for assessing the risk of bias in RCTs [20]. The 2 reviewer authors independently assigned each domain of the Cochrane Collaboration's tool of each individual study to 1 of 3 categories: low, high, or unclear risk of bias. For each study, we created a risk-of-bias table.

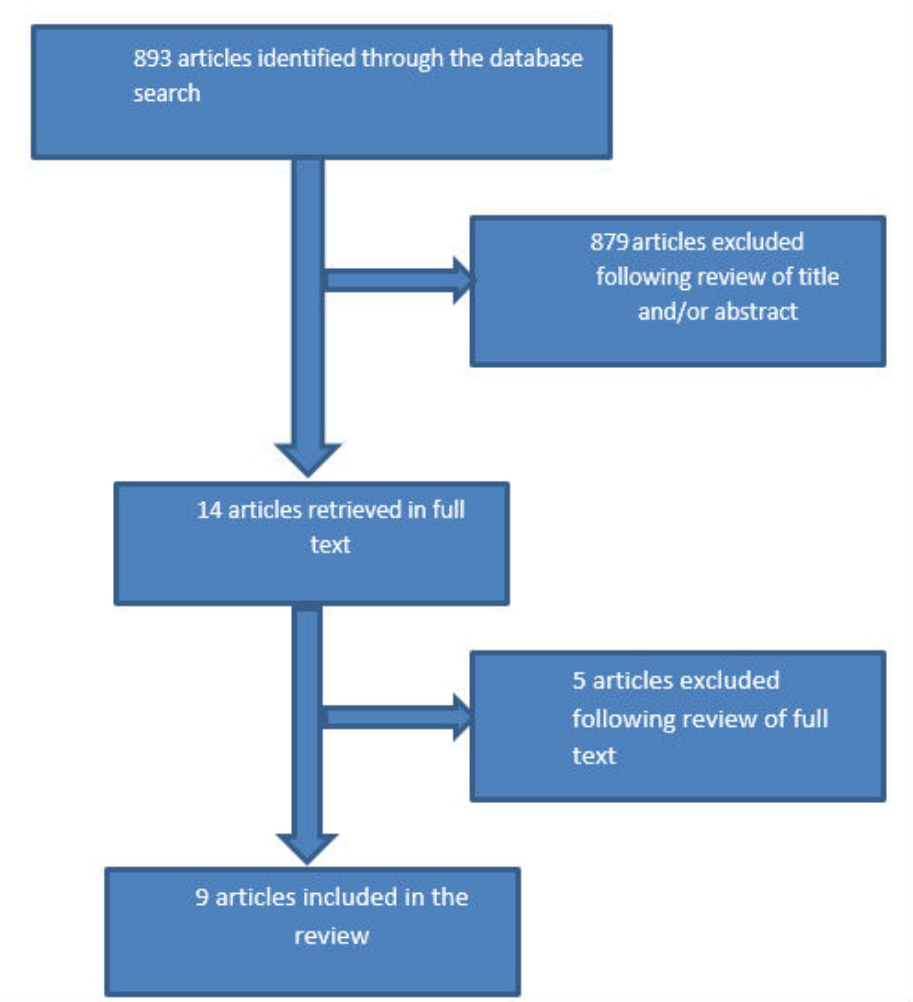
We performed descriptive analyses of the data and summarized the findings from these studies, with emphasis on statistical results reported in RCTs. Differences between groups were highlighted when these results were available. Outcomes were organized into disease-specific clinical outcomes of the intervention. Where available, the usability, feasibility, and acceptability of the intervention were described.

Results

Summary

In all, we reviewed the title or abstract, or both, of 893 papers as retrieved by the database searches. We retrieved 14 papers in full text and assessed them for eligibility. We excluded 5 papers because they did not meet the study design criteria. A total of 9 papers met all inclusion criteria. Figure 1 illustrates the selection process. Multimedia Appendix 1 sets out the quality of the included studies, based on the risk-of-bias assessment. The area in which all studies were assigned a rating of high risk was "blinding of participants and personnel (performance bias)." Due to the nature of the intervention, participants could not be blinded in any of the studies. In some of the studies the personnel involved were also aware of to which group participants were assigned.

Figure 1. Selection process.



Study Characteristics

Publication years ranged from 2008 to 2014 (Multimedia Appendix 2). Studies were included from 4 geographic regions (Europe, n=3; Oceania, n=2; Asia, n=3; United States, n=1). All included studies were RCTs that assessed the effectiveness of interventions that involved mobile phone- or tablet-assisted self-management programs and either standard care (n=8) or offline self-management programs (n=1) [21]. Of the self-management group, 3 included multiple intervention groups using permutations of an intervention involving an app. Charpentier et al [22] ran 2 intervention groups: 1 group used an app plus quarterly clinical visits, and the second used an app plus visits every 2 weeks. Quinn et al [23] ran 3 intervention groups, all involving the patient using an app but with varying degrees of information generated by the app sent to clinicians. Holmen et al [24] ran 2 interventions, 1 using the app alone plus usual care, and 1 using the app and monthly telephone counselling. All 3 involved the use of an app and a Web portal with differing levels of clinician support.

Interventions to address diabetes mellitus (5/9) were the most common, followed by chronic lung disease (3/9) and cardiovascular disease (1/9) interventions. The duration of the intervention ranged from 6 weeks to 1 year, and final follow-up data ranged from 3 months to 1 year. Sample size ranged from 48 to 288 participants.

Impact on Clinical Outcomes

All 9 studies reported the effect of the intervention apps on disease-specific clinical outcomes (Table 1). Significant differences between groups on the primary outcomes were reported in 6 studies [21-23,25-27]. No significant differences on the primary outcomes were found in 2 studies [24,28,29]. Of the 5 interventions related to improving diabetes management [22-25,27], 2 related to type 1 diabetes [22,25] and 3 related to type 2 diabetes [23,24,27]. Significant improvements in diabetes-specific clinical outcomes (HbA_{1c}) were reported in 4 studies [22,23,25,27], 2 relating to type 1 diabetes [22,25] and 2 relating to type 2 diabetes [23,27]. Only 1 study evaluated apps for cardiovascular diseases [29]. Significant improvement in the key clinical outcome, the 6-minute walking test, was observed within the intervention group, indicating overall improved physical functioning, but not between the 2 intervention groups. Mixed results were observed in chronic lung disease clinical outcomes. Lung function parameters were the primary outcomes of interest. Significant improvements were reported in a study on asthma [26] and a study on chronic obstructive pulmonary disease [21], but 1 study, on asthma [28], did not report a significant change.

Table 1. Effectiveness of interventions using mobile phone and tablet apps on primary clinical outcomes (n=9).

Long-term condition addressed	Significant effect, n	No significant effect, n	Total n
Diabetes mellitus type 1	2	0	2
Diabetes mellitus type 2	2	1	3
Cardiovascular disease	0	1	1
Chronic lung diseases	2	1	3
Total clinical outcome studies	6	3	9

Participants

The average age of participants in the studies ranged from 33.8 years [22] to 72.1 years [21]. The age of the participants in relation to ability to use the technology and engagement was not a focus in any study, and none of the studies excluded older people specifically. Holmen et al [24] reported that users >63 years were significantly more likely than younger patients to be substantial users of the app (P=.045).

The majority of studies focused on those with more severe symptoms and clinical indicators outside of the normal range. Of the studies on diabetes, 4 focused on those with HbA_{1c}levels above the optimal range [22-25], and 2 studies [21,26] focused on people with moderate to severe chronic obstructive pulmonary disease and asthma. The characteristics of the target user group was often the impetus for the development of the app tool or mHealth approach, citing the potential for reduced travel to a medical center and ease of engagement as incentives for self-management for these groups with greater symptom burden or higher risk of burden in the future.

In 1 of the studies [25], participants had to provide the phone (and presumably also the network connection) and in 1 study [28], the participants were required to have a contract with a compatible network at their own cost, though a phone could be provided. Neither study noted any significant barriers in recruitment as a result of this inclusion criterion. In all other studies the equipment and the network setup and costs were covered by the research study.

Interventions

The interventions differed in two main ways: the combination of tools used in the intervention and the level of clinician input (Multimedia Appendix 1). Thus, 2 studies [26,28] used an app only and 3 studies ran 2 or 3 intervention groups, 1 of which involved the use of an app only [22-24]. A total of 3 studies [22,23,26] demonstrated a significant change in symptom management and 2 [24,28] did not. An app plus feedback or contact with participants, either by text (4 studies) or phone conversation (3 studies), was used in 7 interventions. A total of 7 studies also used automatic text messages generated by the app data, but only 2 studies [26,28] used this form of feedback alone. In interventions that involved additional clinician input, only 1 study standardized the enhanced clinical input across the control group and intervention [26]. In 6 studies the level of clinician input and support for the intervention group was enhanced. In these studies, the effect of using an app cannot be isolated, and the results cannot be interpreted in relation to increased clinician input or support and the use of an app.

Of the 5 studies that explored the impact of using an app only as the intervention tool on clinical outcomes [22-24,26,28], 3 studies [22,23,26] demonstrated a significant change in symptom management and 2 [24,28] did not. Only 1 study isolated the contribution of the app intervention where clinical care was standardized for the intervention and control group [28]. The study did not report a significant change in asthma symptom control, with symptom control improving marginally in both groups. Similar to the latter study, 4 studies standardized clinical care, with both the control and intervention groups invited to attend one outpatient appointment every 3 months [22-24,26], as per best practice. However, a difference between these 4 studies and the study by Ryan et al [28] was the sharing of app data ahead of the clinic visit [23,26] or the option of sharing the app data during the scheduled clinic visits, in anticipation that the data collected over time would inform the consultation and medical management [22,24]. The impact of this is difficult to quantify. In 2 studies [22,24], sharing data was by the patient's choice, and the number of patients who chose to do so was not reported. In a study involving 2 intervention groups, the effectiveness of the app-only intervention was significant, although the effect size for the app-only intervention group was not as big as that of the intervention group involving an app and teleconsultations [22].

Another study [24] demonstrated an improvement in HbA_{1c}in all 3 groups (a control and 2 intervention groups), although the improvement was not significant for any group. Interestingly, HbA_{1c}decreased more in the app-only group (0.31) than in the control group (0.16) and in the intervention group (0.15) that also involved counselling. In a study on asthma [26], the data were sent to the clinician ahead of the patient consultation, but no data were reported on whether the clinician used this. Only 1 study [23] specifically sought to explore whether sending clinicians' data improved clinical outcomes. All 4 groups continued with usual care (a review with their primary health care provider every 3 months); however, intervention group 1 participants could choose to share their app-generated data with their provider. For intervention group 2 participants, the clinician was sent unanalyzed app-generated data before the scheduled appointment. For intervention group 3 participants, the clinicians were sent analyzed app-generated data ahead of the scheduled appointment. What is not known is how many patients in intervention group 1 shared their data with their clinicians, nor how many clinicians viewed and used the data for participants in groups 2 and 3 in their consultations. The outcome data are mixed, with a decrease in HbA_{1c}noted for all 4 groups, and significant differences noted between the control group and

intervention groups 1 (*P*=.027) and 3 (*P*=.001), but not for intervention group 2 (*P*=.40).

Safety Mechanisms

Only 3 studies [22,27,28] reported an inbuilt safety mechanism for the app intervention, whereby a reading outside of the normal range and considered aberrant enough according to the inbuilt logarithm would trigger an alert. In 1 study [28], the alert would be followed up by the asthma nurse linked with the study. In another [27], an email was generated and sent to the principal investigator and the research nurse. A further study noted the parameters that would be considered abnormal but did not state when or how readings outside of the normal range would be followed up or to whom they would be sent [22]. No study described whether cover included an out-of-hours service; 1 study reported that the data did not trigger any alerts [27]; and the other 2 did not comment on the need for follow-up due to abnormal readings [22,28].

Importantly, no study reported an increase in the number of adverse events or need for additional hospital visits or medical care as a result of participating in the interventions.

Training in the Use of Technology

A total of 5 studies [24,26-29] described training participants in the use of the equipment and input of data. This ranged from distance support [28] with a follow-up 1 week later, to face-to-face support [24,29], and face-to-face support and a 2-week trial as to whether potential participants could use the technology before they were included in the study [27]. Training was mentioned by another study [26] but the nature of this was unclear. In 4 studies participants were offered ongoing technological support, if required, by telephone [24,27-29].

Technological Issues

Few technological issues were reported. Of the studies, 7 required participants to enter the data generated by the study equipment into the app (or via the website) and 2 used a wireless or Bluetooth-compatible device to transmit the data automatically without requiring the participant to manually submit the data [22,24]. No study reported on erroneous imputations (by participants) and only 1 study [24] reported errors in the transfer of data due to issues with the Bluetooth pairing required for automatic transmission of data from the glucometer to the app in the mobile phone.

No study reported on the number of calls participants or clinicians made for technological support during the study. However, 1 study did note that some participants who travelled overseas incurred high mobile costs that weren't covered by the research study. The authors noted that this was anticipated, and participants were informed about the different network rates if travelling.

Usability, Feasibility, and Acceptability

Only 1 study sought to explore the usability and feasibility of the app from the participants' perspective [27], and 1 study explored acceptability of the app from the health care provider's perspective [22]. All studies reported on attrition, which provides some indication of the usability, feasibility, and acceptability of the intervention. The attrition rates ranged from 8.75% [22] to 26% [25]. Most studies reported the attrition rate for the control and intervention groups combined. Where the attrition rates were reported for the control and intervention groups separately, of note, in 1 study, the attrition for the intervention group was higher, with the study average at 25.83%, but the intervention group alone was at 28.33% [26]. In another study, the dropout rate in the usual-care group was considerably higher (36.49% vs 13.21%) [29]. The usual-care group were enrolled in a cardiac rehabilitation program. In a third study [23] the attrition rate was as high as 31.82% in 1 of the intervention groups (control group, 21.45%). The average attrition rate across all 4 groups was 25.6%. No study noted any differences by demographics or clinical parameters between those who dropped out and completers.

A total of 3 studies reported on the number of participants who dropped out specifically because of the technology or because the frequency of input was too burdensome (6/17 [26], 3/7 [29], and 2/3 people [27]).

Only 1 study reported that potential participants were excluded because they could not use the technology [27]. In this study, 12 people were specially excluded before random allocation because they were experiencing difficulties using the devices and sending data. This equated to 18% of the group initially recruited into the trial.

A total of 5 studies explored engagement with the intervention over time and described these findings as a proxy for the usability, feasibility, and acceptability of the intervention [22,24-26,29]. In general, the studies found that the app or intervention was usable, feasible, and acceptable to users. The frequency of data entry was noted to decrease over time in 2 studies [24,25], and there was no significant relationship between level of engagement and change in HbA_{1c}. Another 2 studies reported adherence at the end point of the study as evidence of acceptability (76.7% of the control group and 71.7% of the intervention group [26]; 46.67% of the control group and 80% of the intervention group [29]). A further study asked participants whether they wanted to continue with the intervention if this were possible at the end of the study, reporting that 67% of participants in the app-only intervention and 75% in the app and teleconsultation intervention confirmed that they would [22]. This study was the only one to report on the health care provider's perspective, reporting that 77% of health care providers were satisfied or very satisfied (total number of health care providers unknown). This study found that the time spent by clinicians engaging with participants across the groups was the same (average 71 minutes) but the time saved by the intervention group that did not include hospital clinic visits was 281 minutes (average travel time to attend 2 clinic visits). The time taken to upload data (wirelessly) was estimated to be 10 seconds per day.

How participants felt about the time the intervention took and how easily they had incorporated data entry into their daily life was explored in 1 study [27]. Of 24 participants in the intervention group, 14 felt that the daily data entry was easily incorporated and 10 did not. The average time spent entering data was 22.5 minutes per day. Compliance rates over time (from baseline to the last 2 weeks of the study) for the

components of the intervention differed, with a 70% compliance rate (daily entry) for morning measurements, 50% compliance for bedtime measurements, and 51.2% compliance related to uploading a photo of a meal.

No study reported on the features of the app that patients or health care providers found useful (eg, automated reminders, text messages with educational and motivational content, increased awareness), but 1 study did report that patients expressed feeling reassured, knowing that their health symptoms were regularly monitored [27]. Although the patients' perspective wasn't reported, 2 studies did find that the time saved by not having to visit the hospital for follow-up as per usual care was considerable [22,29].

Costs

The cost implications of using technology were considered in 2 studies. In 1 study where the participant provided the phone and the network connection and a free app was used [25], the cost to the study of messaging participants was calculated at A$290.93, which equated to A $8.08 per participant (n=36). The intervention also included input from a diabetes educator. The diabetes educator spent on average 3 hours per week reviewing participants' logs and text messaging participants, equating to 5 minutes per participant, per week (72 hours in total over the 6-month period). With the hourly rate of A$28.85, the cost to the study was A$2077.20.

The second study required the participant to provide the phone and cover network costs. The technological support service for participants and nursing cover (the safety mechanism for abnormal readings) were contracted out to the company that developed the software, and this was the only cost incurred (not disclosed).

Discussion

Principal Findings

The evidence presented indicates the potential of apps in improving symptom management through self-management interventions. Of the 9 studies, 6 reported a statistically significant difference in the primary clinical outcome of interest. Where the intervention comprised an app only, 3 studies demonstrated a statistically significant improvement. The interventions differed in two main ways: the combination of tools used in the intervention and the level of clinician input (Multimedia Appendix 1). A total of 2 studies [26,28] used an app only and 3 studies ran 2 or 3 intervention groups, 1 of which involved the use of an app only [22-24]. Symptom management changed significantly in 3 studies [22,23,26] but not in 2 [24,28]. An app plus feedback or contact with participants, either by text (4 studies) or phone conversation (3 studies), was used in 7 studies; 7 studies also used automatic text messages generated by the app data, but only 2 studies [26,28] used this form of feedback alone. In interventions that involved additional clinician input, only 1 study standardized the enhanced clinical input across the control group and intervention [26]. In 6 studies the level of clinician input and support for the intervention group was enhanced. In these studies, the effect of using an app cannot be isolated, and the results cannot be interpreted in relation to increased input or support and the use of an app.

Given the evidence that monitoring alone improves symptom control [30], in our review we sought to understand the contribution an intervention involving an app can make. Separating the effect of monitoring alone was not possible in this review. None of the interventions included a study group that involved telemonitoring only. All of the interventions involved either real-time automated feedback or clinician-initiated feedback based on the data entered. The second issue confounding the findings regarding the contribution of the app to symptom management was the use of additional interventions, in combination with the app, to support symptom management.

A further note of caution relates to the ability to generalize the findings to the clinical setting. All of the studies referred to clinical care during the study period as relating to best practice and may indicate an improvement in the actual usual care received. Aside from 1 study [28] where a nurse was employed to ensure follow-up appointments every 3 months, the clinical follow-up appears to have been undertaken in real-life, usual-care clinical settings, although the prompt for data collection at the 3-month time points may have increased the likelihood of these appointments being scheduled and kept. In 1 study [28] the authors specifically noted that, although the clinical care provided during the study period was set up to be in line with best practice with 3-month follow-up appointments, this was likely an improvement in clinical support for the study participants and that it remains unknown whether the addition of the app could improve clinical outcomes when clinical care is less than optimal.

A total of 7 interventions involved the use of the app and a degree of clinical input or support. It does not appear that the intensity of support necessarily affected the outcome, nor the mode of support (eg, electronic, verbal, or face to face). All of the interventions that involved weekly support did note a significant improvement in symptom control, and interventions with a greater gap in time between contact noted mixed results. However, interventions without additional clinical input between usual clinical visits also showed a significant improvement. More frequent clinical input or engagement does not appear to be vital in effecting change in symptom management.

All studies reported minimal issues related to usability. It appears that apps can be used by those with little technology experience or familiarity. The apps do, however, rely on the active engagement of the user. In this review, we could not discern how frequent the level of engagement between user and app needs to be and for how long in order to effect long-term change in symptom management. The engagement of health care providers in monitoring symptoms and exchanging information with users would be desirable in terms of promoting partnership in care, although, again, how important this is in effecting improvements in symptom management is not clear from the review. The freedom and portability of mobile devices, combined with the advanced capacity to facilitate 2-way communication and collect and analyze data for a real-time response, offer enormous potential to patients and health care

providers. The potential complexity of today's mHealth tools and the mixed evidence on the features that are important and make a difference in their effectiveness indicate a need for a focus on understanding the connection between patient experience, adherence, and health outcomes.

The involvement of end users in the development of apps and also specific groups, such as older people and those from different cultures, was not a strong feature in the interventions reviewed. An iterative design process involving systems and content development and multiple stages of user experience testing is recommended for future apps aimed at similar patient populations [31]. The wider evidence suggests that diverse groups can use apps with sufficient training and provision of support [4,32,33], although the level of support over time in both areas remains relatively unexplored.

Technical problems did not feature highly in the review. The issue of erroneous imputations was not mentioned, and errors in transfer of data was noted in only 1 study. However, these areas have been raised in other studies as requiring regular monitoring and attention [34,35].

Given the multitude of apps available, advice on how to develop a "good" app or assess the "quality" of an app when reviewing the existing apps available has important ethical and legal issues for both research and clinical practice. The market has low entry barriers, and ease of accessibility for users through mobile phones and tablets makes this an attractive area for both private and professional areas of application. In health care, when mobile smart devices are used in combination with add-ons that are connected either directly or via wireless technology—for example, blood glucose monitors—manufacturers are required to conform to the laws and regulations that are in place for medical devices, although, depending on the jurisdiction, apps may or may not be well adapted to the specifics of mobile devices. Regulation usually encompasses an app running on a smart device and, for the health professional and patient, some reassurance as to the level of trustworthiness. Stand-alone smart devices and the apps running on them may pose a significant threat to a patient's safety and privacy if the necessary safety measures are not observed. Before recommending or developing an app, it is vital that the functionality be thoroughly tested with respect to the potential for miscalculations, erroneous or incomplete content, technical deficiencies, and other usage restrictions. In addition, assessing the validity of the information and advice regarding symptom management is vital, as well as safety mechanisms relating to when to seek urgent medical advice and support.

The issue of adherence over time was reported by some studies in the review and was found to be unrelated to clinical outcome, that is, those who continued to input data and those who were classified as substantial users did not differ from those whose input declined over time. However, the issue of maintaining engagement over the course of a study and beyond into everyday life is a major consideration and highly likely to affect long-term symptom control. There is a growing understanding of barriers to adherence and ways to overcome them. The development of mAdherence tools to explore barriers to maintaining engagement is growing and will be important in the development of mHealth interventions. In this review we noted differences in patient-provider communication and in the use of targeted motivational messages, but we were not able to qualify their impact.

The issue of enhancing adherence goes beyond maintaining engagement of patients using apps. The development of mHealth tools for chronic disease management could unintentionally increase health disparities in access to technology. Vulnerable, hard-to-reach, or otherwise high-risk patient populations run the risk of exclusion. Where mHealth tools have the potential to engage patients who are less inclined to use traditional health services, mHealth tools offers a way to address barriers to care and reduce health disparities. It is important that future studies specifically build on these areas and reduce the risk of generating a range of interventions largely unused by those who could benefit the most.

Few of the studies we reviewed discussed the issue of cost. The majority of the studies provided the devices to study participants. When implemented at scale, interventions that use patients' existing mobile devices rather than relying on gifted devices will go further toward explaining feasibility and improving clinical outcomes. Rigorous cost-effectiveness analyses will be necessary to demonstrate not only the health impact, but also the value of investing in these innovations.

In addition to cost barriers, other potential barriers for consideration are language and literacy barriers, as well as availability and connectivity issues. Perhaps most critically, if adherence to chronic disease management is not encouraged and actively practiced, it is unlikely that mHealth tools, which are communication platforms and delivery mechanisms, not solutions in and of themselves, will be effective. Conditions such as capability, opportunity, and motivation are essential to behavior change [36]. For example, in diabetes, to improve HbA_{1c}, management of healthy eating, physical activity, and adherence to medication are all important, yet we know little about how we can support and effectively motivate a person in all 3 areas through an app. This again points to the need in future research for the involvement of users as part of the team when developing interventions.

Limitations

There are limitations to this systematic review. A meta-analysis was not possible due to the heterogeneity of the study designs. We did not include non-English literature. The diversity of study objectives, designs, and outcomes made clear comparisons difficult, and the quality of evidence was variable.

Our review expands the evidence base by extending the definition of app interventions to include interventions integrating apps, by assessing both clinical and self-management outcomes, and by contributing to the emerging literature regarding mHealth feasibility, usability, and acceptability.

Conclusion

The use of apps in mHealth has the potential to improve health outcomes among those living with chronic diseases through enhanced symptom control. Further evaluation of apps used in mHealth, and more widely in eHealth, will be valuable. Research

that involves populations traditionally marginalized and research into how these tools can help to overcome barriers to chronic disease management will be especially relevant. Further innovation, optimization, and rigorous research around the potential of apps in mHealth technology will move the field toward the reality of improved health care delivery and outcomes.

Acknowledgments

Both authors contributed to all stages of the review, including conception of the review question, quality appraisal, and synthesis.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Risk of bias assessment for included studies.

[PDF File (Adobe PDF File), 57KB - jmir_v18i5e97_app1.pdf]

Multimedia Appendix 2

Sample characteristics and intervention.

[PDF File (Adobe PDF File), 70KB - jmir_v18i5e97_app2.pdf]

References

1. World Health Organization Media Centre. Diabetes. Geneva: WHO; 2015 Jan. URL: http://www.who.int/mediacentre/factsheets/fs312/en/ [accessed 2015-06-27] [WebCite Cache ID 6ZaZNTIUz]
2. Wilkinson A, Whitehead L, Ritchie L. Factors influencing the ability to self-manage diabetes for adults living with type 1 or 2 diabetes. Int J Nurs Stud 2014 Jan;51(1):111-122. [doi: 10.1016/j.ijnurstu.2013.01.006] [Medline: 23473390]
3. Battersby M, Lawn S, Pols R. Conceptualisation of self-management. In: Kralik D, Paterson B, Coates V, editors. Translating Chronic Illness Research Into Practice. Hoboken, NJ: Wiley-Blackwell; 2010:115-201.
4. Hamine S, Gerth-Guyette E, Faulx D, Green BB, Ginsburg AS. Impact of mHealth chronic disease management on treatment adherence and patient outcomes: a systematic review. J Med Internet Res 2015;17(2):e52 [FREE Full text] [doi: 10.2196/jmir.3951] [Medline: 25803266]
5. de Jongh T, Gurol-Urganci I, Vodopivec-Jamsek V, Car J, Atun R. Mobile phone messaging for facilitating self-management of long-term illnesses. Cochrane Database Syst Rev 2012;12:CD007459. [doi: 10.1002/14651858.CD007459.pub2] [Medline: 23235644]
6. Becker S, Miron-Shatz T, Schumacher N, Krocza J, Diamantidis C, Albrecht U. mHealth 2.0: experiences, possibilities, and perspectives. JMIR Mhealth Uhealth 2014 May;2(2):e24 [FREE Full text] [doi: 10.2196/mhealth.3328] [Medline: 25099752]
7. European Commission. Green Paper on Mobile Health ("mHealth"). Brussels: The Commission; 2015 Feb 23. URL: http://ec.europa.eu/digital-agenda/en/news/green-paper-mobile-health-mhealth [accessed 2015-06-27] [WebCite Cache ID 6ZaZo48ZP]
8. Ofcom. Smartphones Are More Popular Than Laptops Among DE Households. London, UK: Ofcom URL: http://stakeholders.ofcom.org.uk/market-data-research/market-data/communications-market-reports/cmr14/internet-web/uk-4.30 [accessed 2015-06-27] [WebCite Cache ID 6ZaZwYKHX]
9. Ofcom. Among Internet Users in Our European Comparator Countries, Personal Use of Tablets Was Highest in Spain and Italy. London, UK: Ofcom URL: http://stakeholders.ofcom.org.uk/market-data-research/market-data/communications-market-reports/cmr14/international/icmr-5.15 [accessed 2015-06-27] [WebCite Cache ID 6ZaaXpiQ0]
10. Smith A, Page D. U.S. Smartphone Use in 2015. Washington, DC: Pew Internet & American Life Project; 2015 Apr 01. URL: http://www.pewinternet.org/files/2015/03/PI_Smartphones_0401151.pdf [accessed 2016-04-19] [WebCite Cache ID 6gt2P9Hi5]
11. Zickuhr K. Tablet ownership 2013. Washington, DC: Pew Internet & American Life Project; 2013 Jun 10. URL: http://www.pewinternet.org/files/old-media//Files/Reports/2013/PIP_Tablet%20ownership%202013.pdf [accessed 2016-04-19] [WebCite Cache ID 6gt2TobPq]
12. MobiThinking. Global Mobile Statistics 2014 Part A: Mobile Subscribers; Handset Market Share; Mobile Operators. Dublin, Ireland: Afilias Technologies Ltd; 2014 May 16. URL: http://mobiforge.com/research-analysis/global-mobile-statistics-2014-part-a-mobile-subscribers-handset-market-share-mobile-operators [accessed 2015-06-27] [WebCite Cache ID 6Zab3FiII]

13. Lewis TL, Boissaud-Cooke MA, Aungst TD, Eysenbach G. Consensus on use of the term "App" versus "Application" for reporting of mHealth research. J Med Internet Res 2014;16(7):e174; discussion e174 [FREE Full text] [doi: 10.2196/jmir.3460] [Medline: 25033233]
14. Lozano R, Naghavi M, Foreman K, et al. Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet 2012 Dec 15;380(9859):2095-2128. [doi: 10.1016/S0140-6736(12)61728-0] [Medline: 23245604]
15. Marcano Belisario JS, Huckvale K, Greenfield G, Car J, Gunn LH. Smartphone and tablet self management apps for asthma. Cochrane Database Syst Rev 2013;11:CD010013. [doi: 10.1002/14651858.CD010013.pub2] [Medline: 24282112]
16. Terry M. Medical apps for smartphones. Telemed J E Health 2010;16(1):17-22. [doi: 10.1089/tmj.2010.9999] [Medline: 20070172]
17. Zeldes N. History of Computing: The First PDA.: N Zeldes; 2005. URL: http://www.nzeldes.com/HOC/Newton.htm [accessed 2015-06-27] [WebCite Cache ID 6Zae0GUux]
18. Ottawa Hospital Research Institute, University of Oxford. PRISMA: Transparent Reporting of Systematic Reviews and Meta-analyses.: PRISMA; 2015. URL: http://prisma-statement.org/ [accessed 2016-04-19] [WebCite Cache ID 6gt3VO6Wi]
19. Lahdensuo A. Guided self management of asthma--how to do it. BMJ 1999 Sep 18;319(7212):759-760 [FREE Full text] [Medline: 10488007]
20. Higgins J, Green S. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0. 2011. URL: http://handbook.cochrane.org/ [accessed 2015-06-27] [WebCite Cache ID 6Zab5Qh9M]
21. Liu W, Wang C, Lin H, Lin S, Lee K, Lo Y, et al. Efficacy of a cell phone-based exercise programme for COPD. Eur Respir J 2008 Sep;32(3):651-659 [FREE Full text] [doi: 10.1183/09031936.00104407] [Medline: 18508824]
22. Charpentier G, Benhamou P, Dardari D, Clergeot A, Franc S, Schaepelynck-Belicar P, et al. The Diabeo software enabling individualized insulin dose adjustments combined with telemedicine support improves HbA1c in poorly controlled type 1 diabetic patients: a 6-month, randomized, open-label, parallel-group, multicenter trial (TeleDiab 1 Study). Diabetes Care 2011 Mar;34(3):533-539 [FREE Full text] [doi: 10.2337/dc10-1259] [Medline: 21266648]
23. Quinn CC, Shardell MD, Terrin ML, Barr EA, Ballew SH, Gruber-Baldini AL. Cluster-randomized trial of a mobile phone personalized behavioral intervention for blood glucose control. Diabetes Care 2011 Sep;34(9):1934-1942 [FREE Full text] [doi: 10.2337/dc11-0366] [Medline: 21788632]
24. Holmen H, Torbjørnsen A, Wahl AK, Jenum AK, Småstuen MC, Arsand E, et al. A mobile health intervention for self-management and lifestyle change for persons with type 2 diabetes, part 2: one-uear results from the Norwegian randomized controlled trial RENEWING HEALTH. JMIR Mhealth Uhealth 2014;2(4):e57 [FREE Full text] [doi: 10.2196/mhealth.3882] [Medline: 25499872]
25. Kirwan M, Vandelanotte C, Fenning A, Duncan MJ. Diabetes self-management smartphone application for adults with type 1 diabetes: randomized controlled trial. J Med Internet Res 2013;15(11):e235 [FREE Full text] [doi: 10.2196/jmir.2588] [Medline: 24225149]
26. Liu W, Huang C, Wang C, Lee K, Lin S, Kuo H. A mobile telephone-based interactive self-care system improves asthma control. Eur Respir J 2011 Feb;37(2):310-317 [FREE Full text] [doi: 10.1183/09031936.00000810] [Medline: 20562122]
27. Waki K, Fujita H, Uchimura Y, Omae K, Aramaki E, Kato S, et al. DialBetics: a novel smartphone-based self-management support system for type 2 diabetes patients. J Diabetes Sci Technol 2014 Mar 13;8(2):209-215. [doi: 10.1177/1932296814526495] [Medline: 24876569]
28. Ryan D, Price D, Musgrave SD, Malhotra S, Lee AJ, Ayansina D, et al. Clinical and cost effectiveness of mobile phone supported self monitoring of asthma: multicentre randomised controlled trial. BMJ 2012;344:e1756 [FREE Full text] [Medline: 22446569]
29. Varnfield M, Karunanithi M, Lee C, Honeyman E, Arnold D, Ding H, et al. Smartphone-based home care model improved use of cardiac rehabilitation in postmyocardial infarction patients: results from a randomised controlled trial. Heart 2014 Nov;100(22):1770-1779 [FREE Full text] [doi: 10.1136/heartjnl-2014-305783] [Medline: 24973083]
30. American Diabetes Association. Glycemic targets. Diabetes Care 2015 Jan;38 Suppl:S33-S40. [doi: 10.2337/dc15-S009] [Medline: 25537705]
31. Osborn CY, Mulvaney SA. Development and feasibility of a text messaging and interactive voice response intervention for low-income, diverse adults with type 2 diabetes mellitus. J Diabetes Sci Technol 2013 May;7(3):612-622 [FREE Full text] [Medline: 23759393]
32. Lim S, Kang SM, Shin H, Lee HJ, Won YJ, Yu SH, et al. Improved glycemic control without hypoglycemia in elderly diabetic patients using the ubiquitous healthcare service, a new medical information system. Diabetes Care 2011 Feb;34(2):308-313 [FREE Full text] [doi: 10.2337/dc10-1447] [Medline: 21270188]
33. Liang X, Wang Q, Yang X, Cao J, Chen J, Mo X, et al. Effect of mobile phone intervention for diabetes on glycaemic control: a meta-analysis. Diabet Med 2011 Apr;28(4):455-463. [doi: 10.1111/j.1464-5491.2010.03180.x] [Medline: 21392066]
34. Chomutare T, Fernandez-Luque L, Arsand E, Hartvigsen G. Features of mobile diabetes applications: review of the literature and analysis of current applications compared against evidence-based guidelines. J Med Internet Res 2011 Sep;13(3):e65 [FREE Full text] [doi: 10.2196/jmir.1874] [Medline: 21979293]

35. El-Gayar O, Timsina P, Nawar N, Eid W. A systematic review of IT for diabetes self-management: are we there yet? Int J Med Inform 2013 Aug;82(8):637-652. [doi: 10.1016/j.ijmedinf.2013.05.006] [Medline: 23792137]
36. Michie S, van Stralen M, 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]

Abbreviations

HbA1c: glycated hemoglobin
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RCT: randomized controlled trial

Edited by G Eysenbach; submitted 01.07.15; peer-reviewed by P Gee, G Egger; comments to author 31.01.16; revised version received 16.02.16; accepted 20.02.16; published 16.05.16.

Please cite as:
Whitehead L, Seaton P
The Effectiveness of Self-Management Mobile Phone and Tablet Apps in Long-term Condition Management: A Systematic Review
J Med Internet Res 2016;18(5):e97
URL: http://www.jmir.org/2016/5/e97/
doi:10.2196/jmir.4883
PMID:27185295

Original Paper

Potential Reach of mHealth Versus Traditional Mass Media for Prevention of Chronic Diseases: Evidence From a Nationally Representative Survey in a Middle-Income Country in Africa

Maryam Yepes[1,2], RD, PhD; Jürgen Maurer[2], PhD; Barathi Viswanathan[3], RN; Jude Gedeon[3], MD; Pascal Bovet[1,3], MPH, MD

[1]University Institute of Social and Preventive Medicine (IUMSP), Department of Biology and Medicine, Lausanne University Hospital, University of Lausanne, Lausanne, Switzerland

[2]Departments of Econometrics and Political Economics, University of Lausanne, Lausanne, Switzerland

[3]Ministry of Health, Public Health Department, Victoria, Seychelles

Corresponding Author:
Pascal Bovet, MPH, MD
University Institute of Social and Preventive Medicine (IUMSP)
Route de la Corniche 10
1010 Lausanne,
Switzerland
Phone: +41 21 314 7272
Fax: + 41 21 314 737
Email: pascal.bovet@chuv.ch

Abstract

Background: Public radio and television announcements have a long tradition in public health education. With the global rise of computer and mobile device ownership, short message service (SMS) and email-based health services (mHealth) are promising new tools for health promotion.

Objective: Our objectives were to examine 1) self-reported exposure to programs related to noncommunicable diseases (NCDs) on national public television and radio during the 12 months preceding the survey (2013–2014), 2) current ownership of a mobile phone, smartphone, computer, or tablet, and use of the Internet, and 3) willingness of individuals to receive SMS or emails with information on health, with a focus on distribution of these variables across different demographic, socioeconomic status (SES), and NCD risk groups.

Methods: We obtained data in a population survey of 1240 participants aged 25–64 years conducted in 2013–2014 in the Seychelles, a rapidly developing small island state in the African region. We administered a structured questionnaire and measured NCD risk factors. Univariate and multivariate analyses explored the relationships between outcomes and sociodemographic variables.

Results: Of 1240 participants, 1037 (83.62%) reported exposure to NCD-related programs on public television, while a lower proportion of 740 adults (59.67%), reported exposure via public radio (P <.001). Exposure to NCD-related programs on public television was associated with older age (P <.001) and female sex (P <.001), but not with SES, while exposure to NCD-related programs on public radio was associated with older age (P <.001) and lower SES (P <.001). A total of 1156 (93.22%) owned a mobile phone and ownership was positively associated with female sex (P <.001), younger age (P <.001), and higher SES (P <.001). Only 396 adults (31.93%) owned a smartphone and 244 adults (19.67%) used their smartphone to access the Internet. A total of 1048 adults (84.51%) reported willingness to receive health-related SMS, which was positively associated with female sex (P <.001), younger age (P <.001), and higher SES (P <.001). Controlling for SES, exposure to NCD-related programs on public television or radio and willingness to receive health-related SMS were not independently associated with a person's NCD risk.

Conclusions: Broadcasting health programs through traditional mass media (national public radio and television) reached the majority of the population under study, including older adults and those in lower socioeconomic groups. With a high penetration of mobile phones and willingness to receive health-related SMS, mHealth presents an opportunity for health programs, especially when targeted SMS messages are intended for younger adults and those in higher socioeconomic groups. By contrast, due to reduced Internet access, email-based programs had a more limited reach for health promotion programs. These findings emphasize

the different reach of interventions using SMS or email versus traditional mass media, according to demographic and socioeconomic categories, for health education programs in a developing country.

(J Med Internet Res 2016;18(5):e114) doi:10.2196/jmir.5592

KEYWORDS

digital divide; mHealth; eHealth; mass media; mobile phone; noncommunicable diseases; short message service; email; internet access; developing countries; low- and middle-income countries; Africa

Introduction

Evidence suggests that prevention and control of noncommunicable diseases (NCDs) is one of the major health challenges of the 21st century, including in low- and middle-income countries [1,2]. Multifaceted health education strategies combined with multisectoral policies aimed at promoting healthy behaviors are needed to reduce the burden of NCDs [1-4]. Before planning new NCD-related interventions to raise health awareness in the population, it is important to assess the reach of the existing mass media campaigns and to evaluate the potential audience of the planned additional interventions [4]. In this study, we assessed the reach of health education programs on national public television and radio in the Seychelles and the potential reach of short message service (SMS) or email-based interventions (mHealth) according to age, sex, socioeconomic status (SES), and health risk groups.

Health education through traditional mass media such as television and radio, when provided at a sufficiently high frequency, can promote healthy behaviors [5,6]. Advantages of health education campaigns through traditional mass media include a wide audience reach, an easily augmentable frequency of delivery, a high degree of control over content, and a relatively low cost per person exposed [5]. On the other hand, limitations include difficulties in capturing audiences' attention in an increasingly cluttered media environment, the 1-way flow of information from providers to consumers, and a limited ability to offer target-specific messages to pre-identified audiences [6].

To outweigh some of the limitations of health education programs based on traditional mass media, a growing number of SMS or email-based (mHealth) interventions have been used in both developed and developing countries. The steady rise in ownership of mobile phone and other digital communication technologies has facilitated mobile-based interventions, including the provision of health information through mobile messaging and emails [7]. For example, the use of mHealth was highlighted as a key strategy to combat NCDs in developing countries at the 2011 United Nations high-level meeting on NCDs [1]. Pursing this strategy, telecommunication agencies and the World Health Organization (WHO) launched the initiative "Be He@lthy Be Mobile" in 2015, which aims to leverage mobile technology, in particular text messaging and related apps, to help combat the growing global burden of NCDs [8].

In recent decades, there has been a considerable rise in mobile phone ownership in low- and middle-income countries [9]. Based on data from the International Telecommunication Union [9], Figure 1 highlights the growth of mobile phone subscriptions in 48 upper middle-income countries. Figure 1 illustrates that the rate of mobile subscriptions per 100 inhabitants in 50% of upper middle-income countries increased more than 20-fold between 2000 and 2014. The Seychelles—our study site—consistently ranked in the top 25th percentile of upper middle-income countries throughout this period, reaching 160 mobile subscriptions per 100 inhabitants in 2014 [9]. By 2010, the penetration of mobile phones in the Seychelles had surpassed household ownership rates of landline phones (49%), radio (87%), and television (95%) [10].

Mobile-based interventions are increasingly used in health promotion campaigns providing target-specific SMS that encourage specific behavior changes such as increased fruit and vegetable consumption [11], smoking cessation [12], and adoption of healthy lifestyles [13]. Furthermore, mobile-based approaches have also been increasingly applied to address various aspects of disease prevention such as appointment keeping, medication adherence, medical test results delivery, remote diagnosis, data collection, access to health records, disease tracking, and medical response in emergency situations [7,14-18].

SMS- or email-based programs offer several advantages over traditional mass media for health promotion and disease prevention, as they provide opportunities for interactive 2-way communication [15] and target specific, tailored behavior change communication [16,19]. Such interventions offer the opportunity for dissemination of automated, timely, and target-specific messages, which can be designed to complement or mirror in-person counselling [20,21]. For example, messages can offer tailored advice, behavior tracking, goal setting, encouragement, or personal feedback in different stages of behavior change [22-24]. Many theories focus on the need for health messages to offer predisposing, reinforcing, and enabling components of effective health interventions [25]. Several emerging behavior theories suggest that SMS- or email-based interventions can provide timely health messages that can match the level of an individual's motivation and his or her ability to act, and therefore facilitate behavior change [26]. While many studies have shown that mobile phone strategies, using either voice or SMS messaging, can encourage behavior change by increasing patient self-efficacy and assisting in chronic disease management [15,20,24,27-29], further research on SMS-based interventions is required to design effective public health interventions using mobile technology.

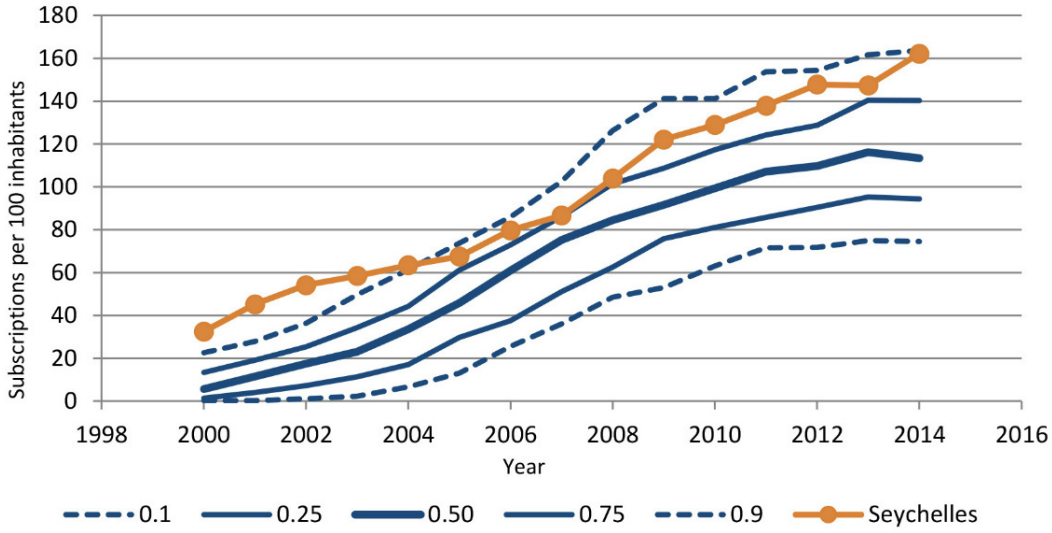
A major constraint in the success of any mHealth initiative is its ability to reach the target population, as well as its adoption, acceptance, and utility from the users' perspective [20,30]. Numerous studies have highlighted the potential inequity in access to technology-based services resulting from differences

in adoption of new technologies such as mobile ownership or Internet access, a concept known as the digital divide [31-33]. Socioeconomic indicators such as sex, education, and income are a few of the many determinants of this digital divide. While the rapid rise in mobile ownership in low- and middle-income countries is reducing the socioeconomic gap in mobile ownership, there remain important differences in adoption of modern technology, particularly in developing countries [31]. Despite this existing gap, several mobile-based initiatives have been successfully implemented in African countries providing health care services to remote areas [24,27-29,34-37].

To date, mHealth interventions in Africa have mostly focused on human immunodeficiency virus/AIDS, malaria, and maternal and child health [20]. However, evidence in support of these approaches in addressing the burden of NCDs in African countries is growing, mainly aimed at improving patient-provider communication [17]. Recent studies have highlighted the benefits of mHealth interventions for cancer care in rural Cameroon [34], Zambia [35], and Nigeria [36] and diabetes care in Western Kenya [37]. However, a continued challenge with health education interventions is to ensure that exposure to health messages extends to all persons targeted, whether a particular group (eg, persons at risk of NCDs) or the entire population, including persons of low SES or otherwise other sociodemographic characteristics that make them less susceptible to be exposed to a particular mass medium.

Despite the overall rise in mobile penetration in developing countries, it is important to assess the penetration of mobile phones and other electronic devices across different demographic, SES, and health risk categories in order to adequately design SMS- or email-based health education interventions. Using data from a national health survey in 2013–2014 in the Seychelles, we assessed 1) self-reported exposure to programs related to NCDs on national public television and radio during the 12 months preceding the survey, 2) current ownership of mobile phones, smartphones, tablets, or computers and Internet access, and 3) willingness of individuals to receive emails or SMS related to information on NCDs, with a focus on the distribution of these variables across different demographic, SES, and NCD risk groups.

Figure 1. Mobile phone subscriptions per 100 inhabitants in percentiles for 48 upper middle-income countries between 2000 and 2014.



Methods

We conducted a nationally representative survey (Seychelles Heart Study IV) in 2013–2014 in the Republic of Seychelles, a rapidly developing small island state in the Indian Ocean, east of Kenya. The survey followed the WHO STEPwise approach to surveillance and was approved by the Ministry of Health of the Seychelles following a technical and ethical review [38,39]. The eligible population included a sex- and age-stratified sample of all adults aged 25-64 years of the 3 main islands, based on computerized data of 2010 national population census, which was thereafter regularly updated by civil authorities. The final study sample consisted of 1240 participants, reflecting a participation rate of 73%.

A structured questionnaire was administered face-to-face by trained survey officers. The questionnaire assessed, among other information, sociodemographic characteristics (age, sex, income, education, and occupation), self-reported health behaviors (including physical activity, smoking, and alcohol habits), exposure to programs related to NCDs on national public television or radio in the past 12 months, ownership and use of selected information and communication technology (mobile phone, computer, tablet, and access to email and Internet), and willingness to receive emails or SMS with information related to health.

Body weight and height were measured and body mass index was calculated as weight (in kg) divided by squared height (in meters). Blood pressure was measured and the average of 3 readings was considered. High blood pressure was defined as blood pressure ≥140/90 mmHg or receiving blood pressure treatment. Elevated blood glucose was defined as plasma glucose >6.1 mmol/L or taking antidiabetic medication. Persons at risk of developing NCD included those who were smoking, drank excessive amounts of alcohol, were not physically active, had high body mass index, or had high blood pressure, high

cholesterol, or diabetes. Multimedia Appendix 1 lists the variables used in this study and their definitions.

We examined the distribution of outcomes across demographic, SES, and NCD risk groups and evaluated group differences using Wald tests. We used logistic regression analysis to assess the associations between socioeconomic variables and our outcome variables (exposure to traditional mass media, ownership of mobile phone and other electronic devices, access to and use of the Internet, and willingness to receive health-related emails and SMS), adjusting for age, sex, and SES. Estimates for crude and adjusted odds ratios were not markedly different and therefore we show only adjusted estimates. Analyses were weighted to reflect the actual population distribution of persons aged 25–64 years. P <.05 was considered significant. Analyses were performed using STATA/SE version 12 software (StataCorp LP).

Results

Table 1 presents the distribution of demographic, SES, and NCD risk categories in the population. Of the 1240 adults, around half had obtained education beyond obligatory education (ie, beyond the age of 15–16 years, and approximately one-third of adults earned more than 8000 Seychelles rupees per month (~US$ 600), in line with data from the Seychelles National Bureau of Statistics [40]. In the sample of 1240 participants aged 25–64 years, the prevalences were 64.32% (n=798) for overweight or obesity, 59.90% (n=743) for hypertension, 21.36% (n=265) for elevated blood glucose, 20.60% (n=255) for smoking, 25.41% (n=315) for heavy alcohol intake, and 23.16% (n=287) for sedentary lifestyle. All proportions are adjusted to the actual age and sex distributions of the population of Seychelles.

Mobile Technology Ownership

Table 2 shows the univariate distribution of ownership of electronic devices and Internet access by sex, age, and SES categories and the multivariate odds ratios for the relationships between the outcomes when adjusted for all the variables displayed in the table. Overall, of the 1240 adults aged 25–64 years, 1156 (93.21%) owned a mobile phone. The distribution differed by sex: 1191 women (96.04%) versus 1120 men (90.32%); age group: 1197 aged 25–34 (96.58%) versus 996 aged 55–64 (80.34%) years; and SES: 1228 persons with university education (99.03%) versus 860 persons who did not complete obligatory school (69.35%). Of the 1240 adults, 396 (31.93%) owned a smartphone, 245 (19.75%) used their smartphone to access the Internet, 678 (54.67%) owned a computer or tablet at home, and 477 (38.46%) had Internet access with a personal computer or tablet. The social patterning of these outcomes was generally comparable with that of mobile phone ownership described above. We observed similar age and SES differences in adjusted models, even if these associations were not always statistically significant.

Exposure to Health Programs on Mass Media

The first 2 columns of Table 3 show the exposure to health programs on NCDs through traditional mass media (limited to health programs from national public radio and television broadcasting company). It is important to note that, in our sample of 1240 adults, 1234 (99.51%) owned a television and 1221 (98.46%) owned a radio. In Table 3 , the reported proportions are univariate, while odds ratios are multivariate, adjusted to other variables. Participants were asked about their exposure to health education programs on national public television or radio on cardiovascular disease, including stroke and heart attack; cardiovascular risk factors such as hypertension, blood cholesterol, or diabetes; and lifestyle habits such as smoking, alcohol drinking, diet, and physical activity. Overall, 1036 of the 1240 participants (83.54%) reported having viewed at least one such program on public television during the past 12 months, while only 740 of the 1240 participants (59.67%) reported having listened to at least one such program on public radio during the same time period.

In multivariate analysis, viewing health-related television programs was higher among women and more mature adults, and there was no clear systematic relationship with SES. Similarly, exposure to health-related programs on public radio was higher among women and older adults, but negatively associated with SES, even though the estimated coefficients were generally not statistically significant.

Willingness to Receive Health Massages by SMS or Email

The last 2 columns of Table 3 present willingness to receive health-related SMS or email messages. For self-reported willingness to receive health-related messages, 1048 of the 1240 (84.51%) participants expressed willingness to receive such information by SMS compared with only 508 participants (40.96%) by email. In univariate analyses, this proportion was significantly associated with female sex: 1096 women (88.39%) versus 1000 men (80.64%); age: 1128 young adults at age 25–34 (90.96%) versus 808 older adults at age 55–64 (65.21%,) years; and SES: 1031 adults with highest (83.14%) versus 634 adults with lowest education categories (51.12%). When adjusted for all explanatory variables at the same time, willingness to receive SMS or emails related to health continued to be positively associated with female sex, younger age, and higher SES, consistent with results of the univariate analyses.

Table 4 examines the proportion of those with NCD risk factors with the likelihood of exposure to NCD-related health programs on national public media and willingness to receive SMS or emails related to health, when unadjusted and adjusted for age, sex, and SES, in addition to all NCD variables. Univariate analysis showed a higher proportion of those with hypertension and diabetes having watched television or radio programs on health in the past 12 months. However, these differences are no longer significant in a multivariate model when adjusted for SES variables. Findings were similar for willingness to receive health-related SMS or email messages. These findings suggest that when individuals' variation in age, sex, income, education, and occupation are taken into account, NCD risk factors (ie, hypertension, elevated blood glucose, overweight, smoking, or high alcohol intake) do not significantly increase or decrease their interest in NCD-related interventions via mass media or mHealth.

Table 1. Characteristics of the study population (Seychelles, N=1240, 2013–2014).

Population characteristics		No.	%
Sociodemographic variables			
Sex			
	Male	621	50.11
	Female	620	49.98
Age (years)			
	25–34	356	28.68
	35–44	362	29.35
	45–54	327	26.36
	55–64	197	15.85
Education			
	Did not complete obligatory	91	7.37
	Completed obligatory	560	45.16
	Vocational	174	14.02
	Polytechnic	344	27.76
	University	72	5.77
Annual income (Seychelles rupees)			
	<3000	106	8.52
	3000–5000	312	25.17
	5000–8000	441	35.53
	8000–15,000	280	22.59
	≥15,000	102	8.20
Occupation			
	Nonqualified laborer	254	20.46
	Semiqualified manual	351	28.28
	Qualified manual	193	15.57
	Semiqualified nonmanual	256	20.64
	Qualified professional	224	18.05
Clinical variables			
Body mass index (kg/m ^{2})			
	<25 (normal weight)	432	34.82
	25–30 (overweight)	424	34.22
	≥30 (obese)	373	30.10
Physical activity			
	Sedentary	287	23.16
	Moderate	780	62.90
	Active	173	13.95
Alcohol consumption			
	None	524	42.22
	Moderate	401	32.37
	Heavy	315	25.41
Smoking status			

Population characteristics		No.	%
	Current smoker	255	20.60
	Ex-smoker	119	9.57
	Never smoked	866	69.84
Hypertension			
	Negative	497	40.10
	Positive	743	59.90
Elevated blood sugar			
	Negative	975	78.66
	Positive	265	21.36

Table 2. Univariate distribution and multivariate odds ratios for ownership of mobile phone, smartphone, computer, or tablet, and Internet access, by sex, age and socioeconomic status (Seychelles, N=1240, 2013–2014).

Population characteristics	Owns a mobile phone		Owns a smartphone		Has access to Internet with smartphone		Owns a personal tablet or computer		Has access to Internet with personal computer or tablet	
	%	aOR[b]	%	aOR	%	aOR	%	aOR	%	aOR
Total	93.2		31.9		19.7		54.3		38.5	
Sex										
Male	90.4	1	28.9	1	18.0	1	51.7	1	38.4	1
Female	96.1	*6.81* [c]	34.9	*1.44*	21.4	*1.49*	56.9	*1.52*	38.6	1.09
P value	<.001		.03		.17		.08		.94	
Age (years)										
25–34	96.6	1	51.8	1	33.7	1	63.6	1	47.8	1
35–44	96.5	0.72	30.8	*0.40*	18.5	*0.43*	62.0	0.93	42.4	0.74
45–54	93.6	0.61	24.8	*0.34*	13.3	*0.33*	49.5	*0.59*	35.1	*0.60*
55–64	80.3	*0.27*	9.90	*0.15*	7.10	*0.22*	31.3	*0.37*	20.1	*0.38*
P value	<.001		<.001		<.001		<.001		<.001	
Education										
Did not complete obligatory	69.4	1	6.10	1	4.70	1	17.4	1	10.1	1
Obligatory	91.8	*1.98*	19.8	1.84	10.00	1.02	38.1	1.31	20.1	0.87
Vocational	95.9	2.54	43.9	*3.85*	28.7	2.48	58.8	*2.66*	30.1	1.82
Polytechnic	99.2	*8.65*	45.8	*3.17*	28.3	2.03	80.0	*3.45*	65.3	2.82
University	99.1	7.70	63.9	*7.42*	51.5	*5.53*	94.5	*6.51*	87.6	*5.64*
P value	<.001		<.001		<.001		<.001		<.001	
Annual income (Seychelles rupees)										
<3000	74.3	1	9.50	1	4.10	1	30.7	1	15.7	1
3000–5000	90.8	*2.62*	26.7	1.87	16.3	2.14	37.0	0.69	20.0	0.61
5000–8000	95.1	*5.45*	30.4	1.95	17.0	1.93	50.4	0.86	34.5	0.89
8000–15,000	98.0	*7.32*	45.1	*2.55*	28.3	2.68	75.0	1.57	59.7	1.45
≥15,000	98.7	*13.8*	41.4	1.93	34.2	*3.45*	92.1	*6.14*	77.9	*3.07*
P value	<.001		<.001		<.001		<.001		<.001	
Occupation										
Nonqualified laborer	83.0	1	13.3	1	5.00	1	22.5	1	6.90	1
Semiqualified manual	93.5	*3.54*	29.6	*1.97*	19.1	*3.31*	43.6	*2.15*	30.2	*4.23*
Qualified manual	92.8	2.25	26.6	1.29	17.7	*2.33*	55.4	*2.47*	34.7	*3.28*
Semiqualified nonmanual	98.6	*4.84*	36.9	*2.14*	22.5	*3.08*	71.0	*4.26*	51.4	*6.72*
Qualified professional	98.4	2.31	54.5	*3.13*	35.4	*3.28*	87.2	*5.53*	75.1	*8.32*
P value	<.001		<.001		<.001		<.001		<.001	

[a]Wald test for univariate differences between the first and last categories of each socioeconomic variable.

[b]Multivariate odds ratios are adjusted (aOR) to all socioeconomic variables (sex, age, education, income, and occupation) with the reference categories identified as 1.

[c]Coefficients with *P* <.05 are represented in italics.

Table 3. Univariate distribution and multivariate odds ratios for exposure to television or radio programs on NCDs during the past 12 months and willingness to receive health-related SMSc or email, by sex, age and socioeconomic status (Seychelles, N=1240, 2013–2014).

Population characteristic	Viewed a program on television		Listened a program on radio		Would like to receive SMS messages		Would like to receive email messages	
	%	aOR[d]	%	aOR	%	aOR	%	aOR
Total	83.6		59.7		84.5		41.0	
Sex								
Male	79.3	1	57.8	1	80.7	1	36.4	1
Female	87.8	*1.92* [e]	61.6	1.08	88.4	*2.73*	45.3	*1.65*
P value	<.001		.2		<.001		.006	
Age (years)								
25–34	70.1	1	43.8	1	91.5	1	49.5	1
35–44	87.2	*2.76*	55.2	1.66	87.1	*0.65*	43.7	*0.61*
45–54	90.0	*3.62*	71.7	3.41	86.3	*0.85*	39.9	*0.54*
55–64	88.9	*3.20*	76.6	3.88	65.2	*0.33*	19.7	*0.26*
P value	<.001		<.001		<.001		<.001	
Education								
Did not complete obligatory	86.0	1	80.0	1	51.1	1	10.3	1
Obligatory	87.1	*1.27*	62.0	0.72	83.0	2.57	21.5	*0.65*
Vocational	74.1	0.84	52.7	0.80	90.1	3.63	32.4	1.04
Polytechnic	81.7	1.09	55.6	0.96	93.4	5.98	67.5	1.56
University	84.7	*1.47*	51.9	0.79	83.2	3.61	93.3	*4.93*
P value	.85		<.001		<.001		<.001	
Annual income (Seychelles rupees)								
<3000	85.5	1	75.7	1	60.1	1	10.0	1
3000–5000	85.3	*1.31*	64.5	0.88	84.5	*2.37*	20.6	*0.92*
5000–8000	82.6	*1.17*	56.7	0.69	87.7	*2.62*	34.7	*1.26*
8000–15,000	82.9	1.22	55.6	0.69	89.3	2.22	63.3	2.29
≥15,000	82.3	0.81	52.0	0.46	83.5	1.99	82.8	5.94
P value	.55		<.001		<.001		<.001	
Occupation								
Nonqualified laborer	85.8	1	72.3	1	69.4	1	3.90	1
Semiqualified manual	83.0	*1.21*	60.2	0.75	88.1	*3.34*	23.6	*5.91*
Qualified manual	77.3	*0.97*	51.3	0.57	82.5	*1.73*	33.7	*6.46*
Semiqualified nonmanual	86.0	*1.06*	54.0	0.55	93.5	*3.08*	59.0	*16.7*
Qualified professional	83.4	0.96	56.7	0.64	87.2	1.30	82.0	25.4
P value	.49		<.001		.001		<.001	

[a]Wald test for univariate differences between the first and last categories of each socioeconomic variable.

[b]NCDs: noncommunicable diseases.

[c]SMS: short message service.

[d]Multivariate odds ratios are adjusted (aOR) to all socioeconomic variables (sex, age, education, income, and occupation) with the reference categories identified as 1.

[e]Coefficients with *P* <.05 are represented in italics.

Table 4. Univariate distribution[a] and multivariate odds ratios for exposure to television or radio programs on NCDs[b] during the past 12 months and willingness to receive health-related SMS[c] or email, by NCD risk factor categories (Seychelles, n=1240, 2013–2014).

Population characteristic	Mobile ownership	Viewed a program on television		Listened a program on radio		Would like to receive SMS messages		Would like to receive email messages	
	%	%	aOR	%	aOR	%	aOR	%	aOR
Smoking									
Smoker	83.3	81.3	1	55.8	1	80.2	1	28.0	1
Ex-smoker	92.2	85.5	0.93	62.6	0.99	84.5	1.31	39.8	1.47
Never smoked	96.2	83.9	0.82	60.4	1.24	85.8	0.87	44.5	0.96
	<.001	.41		.25		.06		<.001	
Alcohol intake									
None	94.5	84.8	1	62.3	1	84.6	1	40.7	1
Moderate	94.8	84.3	1.24	58.6	1.05	85.9	1.13	46.2	0.88
Heavy	89.0	80.3	0.86	56.5	0.91	82.7	0.88	34.5	0.65
.006	.14		.12		.50		0.13		
Physical activity									
Sedentary	93.6	83.9	1	58.6	1	86.4	1	59.0	1
Moderate	93.2	84.8	1.09	60.9	1.03	85.2	0.94	39.8	*0.53* [e]
Active	92.7	77.0	0.81	55.7	0.99	78.3	0.57	19.3	*0.3*
	.72	.11		.58		.05		<.001	
Body mass index									
<25	89.2	82.0	1	53.0	1	83.0	1	38.3	1
25–30	95.7	82.5	0.68	61.1	1.15	84.8	1.01	43.3	0.98
≥ *30*	94.8	86.4	0.7	65.4	1.22	86.0	1.05	41.2	0.98
	.002	.11		<.001		.42		.49	
Hypertension									
Negative	94.9	79.9	1	53.9	1	87.9	1	45.0	1
Positive	90.7	89.0	1.27	68.3	1.05	79.4	0.74	34.8	0.87
	.003	<.001		<.001		<.001		.002	
Elevated sugar									
Negative	94.8	81.8	1	56.7	1	86.9	1	42.5	1
Positive	87.2	90.0	1.36	70.7	1.09	75.8	0.79	34.9	1.23
	<.001	.001		<.001		<.001		.002	

[a]Wald test for univariate differences between the first and last categories of each socioeconomic variable.

[b]NCDs: noncommunicable diseases.

[c]SMS: short message service.

[d]Multivariate odds ratios are adjusted (aOR) to all socioeconomic variables (sex, age, education, income, and occupation) with the reference categories identified as 1.

[e]Coefficients with P <.05 are represented in italics.

Discussion

We found a large exposure to programs related to NCDs on national public television in the adult population aged 25–64 years in the Seychelles. This exposure was especially large among women and older persons, with no significant association with SES. On the other hand, exposure to health programs on radio was lower than on that on television, with higher exposure among persons of lower SES than among higher SES. We found that the majority of adults owned a mobile phone, but fewer owned smartphones, computers, or tablets or had Internet access. The willingness to receive health-related SMS was higher in women, younger adults, and those in higher SES. We also found that willingness to receive health-related SMS was not

independently related to a person's NCD risk. Overall, this study highlights the different reach, according to age, sex, and SES, of health messages on NCDs supplied through public mass media programs versus health messages that would be based on SMS or email. In particular, our findings emphasize the presence of a digital divide according to age, sex, and SES despite the large penetration of mobile phone and other new electronic media in the population. This divide does not mean that modern media are inappropriate for health education programs related to NCDs, but it suggests that interventions based on mobile technology should be carefully designed with regard to specific purposes and audiences. Our findings may have relevance for other countries that are similar to the Seychelles, including other small island developing states or some middle-income countries that have achieved rapid socioeconomic development.

The sociodemographic differences in mobile ownership and willingness to receive health-related SMS or email messages, as found in this study, play a key role in ensuring that mHealth initiatives have an equitable reach among the target populations [30]. While SMS-based interventions have a promising potential for disease prevention and health service delivery, such interventions can also potentially exacerbate health inequalities arising from a digital divide [41]. Our findings are consistent with a social pattern in the uptake of health-related information [42].

In addition to a social digital divide, numerous studies have also suggested a sex divide in mobile technology adoption, with greater mobile access among men than among women in most developing countries [33,43,44]. In contrast, we found that more women than men owned a mobile phone or a smartphone. Women were also more willing to receive health-related SMS and were more likely to watch or listen to NCD-related programs on public television and radio during the year under study. These differences may be small in absolute magnitude but can be important when designing health-related public education programs. This sex difference favoring women over men may be partially explained by the relatively high sex equity in Seychelles, for example, the Seychelles ranked second highest out of 52 African countries for sex equity according to the Ibrahim Index [45].

We also found an age-related digital divide, with modern communication technologies being used more often by younger than by older persons. This age-related digital divide was small for ownership of mobile phones but was larger for smartphone ownership, and access to and use of the Internet. Also, fewer older than younger persons were willing to receive SMS or emails related to health. This is consistent with younger adults and persons of higher SES being prone to adopt new technologies [33,46,47].

We found that persons at higher risk of specific NCD conditions were not more likely to watch or listen to NCD-related public television or radio programs, and were not more willing to receive health-related SMS or emails than were persons at lower NCD risk, when adjusted for demographics and SES. It is possible that persons with NCD risk would have shown more interest in SMS-based health programs if they had been fully aware of the many potential benefits of mHealth in general, such as SMS reminders to attend medical visits or to take medications, or health messages tailored to a person's particular condition.

The finding that the use of mobile communication devices, or willingness to receive health-related SMS or email messages, is higher among younger adults and among persons of higher SES—irrespective of NCD risk—has important public health implications, since NCDs tend to concentrate in persons of older age and in persons of lower SES. The social and age-related digital divides not only stem from different levels of ownership of mobile devices, but also relate to differences in motivation, ability, and skills to use these devices and the related apps. It is likely that the age and socioeconomic digital divides will decrease in the coming generations, which will allow for persons from broader age and SES categories also to benefit from mHealth services. More generally, age, sex, and social differences in the use of mHealth stress the need to carefully design interventions and to address potential equity concerns.

On the other hand, the larger ownership of mobile phones and smartphones, and access to the Internet, as well as the larger acceptance of health-related SMS and emails, among younger adults may be viewed as an advantage when designing SMS-based mHealth initiatives among young adults. In the same line, the frequent use of email and Internet among persons with qualified occupations suggests that this target population could benefit most from Web-based and email health messaging at the work place.

We also found that, despite the overall rise in the use of new media and communication technologies, large proportions of adults continue to listen to and watch health programs on radio and television. This is consistent with a strong commitment from the Seychelles' health authorities to frequently broadcast radio and television programs on health during the past 25 years. We found that exposure to NCD-related programs on television was fairly uniform across age, sex, and SES groups. Hence, these programs may be useful to educate the public about health issues of general interest, such as raising awareness of NCDs in the general public, emphasizing the importance of good medication adherence, explaining the components of a healthy diet, or stressing the importance of regular blood pressure checks. We also found that programs on NCDs on public radio had a larger audience among older persons and among persons of lower SES, which also corresponds to the population subgroups at higher risk of NCD. Hence, there may be some benefit for continued use of traditional media to broadcast health education programs related to NCDs in developing countries such as the Seychelles.

Several limitations of this study need to be highlighted. The age range was limited to 25–64 years, that is, we excluded both children (when healthy behaviors are ingrained) and the elderly (those at higher risk of NCDs). Also, our questions on willingness to receive health-related email or SMS were not designed to assess the whole range of mHealth services. Similarly, questions on exposure to health programs on traditional media did not assess the impact of these programs on behaviors. Finally, while the findings in our study may

possibly extend to a few other rapidly developing small island states or certain upper middle-income countries that have experienced rapid socioeconomic development similar to that in the Seychelles, further research is required to replicate these findings in similar contexts. Strengths of the study include the population-based design, the fairly large sample size, and the assessment of numerous variables related to the use of mobile technology and to the exposure to health programs related to NCDs.

In conclusion, our study offers new evidence on exposure of health-related programs through traditional media and the feasibility and acceptance of mHealth interventions in an upper middle-income country. With a high reach among all groups of the population, national television programs on health appeared to continue to serve as a valuable medium for health promotion and NCD prevention. A large majority of the population owned a mobile phone and were willing to receive health-related SMS messages. However, due to heterogeneous distribution of mobile technology and a digital divide at the time of the survey, mHealth intervention showed the highest potential reach among persons with higher income and education, as well as in younger adults. These finding are important to design and implement mHealth interventions or health programs in the mass media. More generally, this study highlights advantages and disadvantages of traditional mass media versus modern mobile technology for providing health education and the substantial differences in exposure that can occur according to age, sex, and SES. Our findings further emphasize that health education interventions supplied through traditional mass media or through modern mobile technology must be carefully designed in terms of the intended targeted audiences.

Acknowledgments

MY, JM, and PB contributed to the analysis and writing of the manuscript. PB, BV, and JG contributed to study design and data collection, as well as critical review of the manuscript. Grateful thanks are expressed to the survey officers, to all the participants in the survey, and to the Ministry of Health of the Republic of Seychelles for continued support of epidemiological research.

MY's contribution was funded by the Swiss National Fund (PDF MPI_137232); data collection was partially funded by the Ministry of Health, Seychelles; the University Institute of Social and Preventive Medicine, Lausanne, Switzerland; the World Health Organization (Regional Office for Africa); and unconditional grants from the Seychelles Trading Company (STC Ltd) and Seychelles Petroleum (SEYPEC, Ltd).

Conflicts of Interest

None declared.

Multimedia Appendix 1

Study survey questions and response categories.

[PDF File (Adobe PDF File), 159KB - jmir_v18i5e114_app1.pdf]

References

1. World Health Organization. Global Action Plan For The Prevention and Control Of Noncommunicable Diseases. Geneva: WHO Press; 2013. URL: http://apps.who.int/iris/bitstream/10665/94384/1/9789241506236_eng.pdf?ua=1 [accessed 2016-04-28] [WebCite Cache ID 6h725PwkS]
2. Abegunde DO, Mathers CD, Adam T, Ortegon M, Strong K. The burden and costs of chronic diseases in low-income and middle-income countries. Lancet 2007 Dec 08;370(9603):1929-1938 [FREE Full text] [doi: 10.1016/S0140-6736(07)61696-1]
3. Cecchini M, Sassi F, Lauer JA, Lee YY, Guajardo-Barron V, Chisholm D. Tackling of unhealthy diets, physical inactivity, and obesity: health effects and cost-effectiveness. Lancet 2010 Nov 10;376(9754):1775-1784. [doi: 10.1016/S0140-6736(10)61514-0] [Medline: 21074255]
4. Gaziano T, Galea G, Reddy K. Scaling up interventions for chronic disease prevention: the evidence. Lancet 2007 Dec 05;370(9603):1939-1946. [doi: 10.1016/S0140-6736(07)61697-3]
5. Wakefield MA, Loken B, Hornik RC. Use of mass media campaigns to change health behaviour. Lancet 2010 Oct 9;376(9748):1261-1271 [FREE Full text] [doi: 10.1016/S0140-6736(10)60809-4] [Medline: 20933263]
6. Naugle DA, Hornik RC. Systematic review of the effectiveness of mass media interventions for child survival in low- and middle-income countries. J Health Commun 2014;19 Suppl 1:190-215 [FREE Full text] [doi: 10.1080/10810730.2014.918217] [Medline: 25207453]
7. Krishna S, Boren SA, Balas EA. Healthcare via cell phones: a systematic review. Telemed J E Health 2009 Apr;15(3):231-240. [doi: 10.1089/tmj.2008.0099] [Medline: 19382860]
8. International Telecommunication Union. ITU-WHO Mobile Health for Non-Communicable Diseases (NCDs) Initiative. Geneva, Switzerland: ITU; 2015. URL: http://www.itu.int/en/ITU-D/ICT-Applications/Pages/Be_Healthy_intro.aspx [accessed 2016-02-02] [WebCite Cache ID 6f02JQU9I]

9. The World Bank. Maximizing Mobile: Information and Communications for Development. Washington, DC: The World Bank; 2012. URL: http://siteresources.worldbank.org/EXTINFORMATIONANDCOMMUNICATIONANDTECHNOLOGIES/Resources/IC4D-2012-Report.pdf [accessed 2016-02-02] [WebCite Cache ID 6f03HUJ1B]
10. InternationalTelecommunication Union. Statistics: Mobile-Cellular subscriptions. Geneva, Switzerland: ITU; 2015. URL: http://www.itu.int/en/ITU-D/Statistics/Pages/stat/default.aspx [accessed 2016-02-02] [WebCite Cache ID 6f00UzunJ]
11. Silva C, Fassnacht DB, Ali K, Gonçalves S, Conceição E, Vaz A, et al. Promoting health behaviour in Portuguese children via short message service: the efficacy of a text-messaging programme. J Health Psychol 2015 Jun;20(6):806-815. [doi: 10.1177/1359105315577301] [Medline: 26032797]
12. International Telecommunication Union. Costa Rica Ministry of Health Launches mHealth for Smoking Cessation. Geneva, Switzerland: ITU; 2013. URL: http://www.itu.int/en/ITU-D/ICT-Applications/eHEALTH/Be_healthy/Pages/mHealth_CostaRica_smoking.aspx [accessed 2016-02-02] [WebCite Cache ID 6f02dxd7g]
13. Beratarrechea A, Diez-Canseco F, Fernández A, Kanter R, Letona P, Martinez H, et al. [Acceptability of a mobile health based intervention to modify lifestyles in prehypertensive patients in Argentina, Guatemala and Peru: a pilot study]. Rev Peru Med Exp Salud Publica 2015;32(2):221-229 [FREE Full text] [Medline: 26338378]
14. World Health Organization. mHealth: New Horizons for Health Through Mobile Technologies. Geneva, Switzerland: WHO Press; 2011. URL: http://apps.who.int/iris/bitstream/10665/44607/1/9789241564250_eng.pdf [accessed 2016-02-02] [WebCite Cache ID 6f02paBWF]
15. Cole-Lewis H, Kershaw T. Text messaging as a tool for behavior change in disease prevention and management. Epidemiol Rev 2010;32:56-69 [FREE Full text] [doi: 10.1093/epirev/mxq004] [Medline: 20354039]
16. Bastawrous A, Armstrong MJ. Mobile health use in low- and high-income countries: an overview of the peer-reviewed literature. J R Soc Med 2013 Apr;106(4):130-142 [FREE Full text] [doi: 10.1177/0141076812472620] [Medline: 23564897]
17. Bloomfield GS, Vedanthan R, Vasudevan L, Kithei A, Were M, Velazquez EJ. Mobile health for non-communicable diseases in Sub-Saharan Africa: a systematic review of the literature and strategic framework for research. Global Health 2014;10:49 [FREE Full text] [doi: 10.1186/1744-8603-10-49] [Medline: 24927745]
18. Thakkar J, Kurup R, Laba T, Santo K, Thiagalingam A, Rodgers A, et al. Mobile telephone text messaging for medication adherence in chronic disease: a meta-analysis. JAMA Intern Med 2016 Mar 1;176(3):340-349. [doi: 10.1001/jamainternmed.2015.7667] [Medline: 26831740]
19. Déglise D, Suggs LS, Odermatt P. SMS for disease control in developing countries: a systematic review of mobile health applications. J Telemed Telecare 2012 Jul;18(5):273-281. [Medline: 22826375]
20. Gurman TA, Rubin SE, Roess AA. Effectiveness of mHealth behavior change communication interventions in developing countries: a systematic review of the literature. J Health Commun 2012;17 Suppl 1:82-104. [doi: 10.1080/10810730.2011.649160] [Medline: 22548603]
21. Kamis K, Janevic MR, Marinec N, Jantz R, Valverde H, Piette JD. A study of mobile phone use among patients with noncommunicable diseases in La Paz, Bolivia: implications for mHealth research and development. Global Health 2015;11:30 [FREE Full text] [doi: 10.1186/s12992-015-0115-y] [Medline: 26141528]
22. Head KJ, Noar SM, Iannarino NT, Grant Harrington N. Efficacy of text messaging-based interventions for health promotion: a meta-analysis. Soc Sci Med 2013 Nov;97:41-48. [doi: 10.1016/j.socscimed.2013.08.003] [Medline: 24161087]
23. Abroms LC, Whittaker R, Free C, Mendel Van Alstyne, Schindler-Ruwisch JM. Developing and pretesting a text messaging program for health behavior change: recommended steps. JMIR Mhealth Uhealth 2015;3(4):e107 [FREE Full text] [doi: 10.2196/mhealth.4917] [Medline: 26690917]
24. Piette JD, List J, Rana GK, Townsend W, Striplin D, Heisler M. Mobile health devices as tools for worldwide cardiovascular risk reduction and disease management. Circulation 2015 Nov 24;132(21):2012-2027. [doi: 10.1161/CIRCULATIONAHA.114.008723] [Medline: 26596977]
25. Green LW, Kreuter MW. Health Promotion Planning: An Educational and Environmental Approach. Mountain View, CA: Mayfield; 2005.
26. Fogg BJ. A behavior model for persuasive design. California, USA: Claremont; 2009 Presented at: 4th international Conference on Persuasive Technology; April 26-29, 2009; Claremont, CA.
27. Gurol-Urganci I, de Jongh T, Vodopivec-Jamsek V, Atun R, Car J. Mobile phone messaging reminders for attendance at healthcare appointments. Cochrane Database Syst Rev 2013;12:CD007458. [doi: 10.1002/14651858.CD007458.pub3] [Medline: 24310741]
28. de Jongh T, Gurol-Urganci I, Vodopivec-Jamsek V, Car J, Atun R. Mobile phone messaging for facilitating self-management of long-term illnesses. Cochrane Database Syst Rev 2012;12:CD007459. [doi: 10.1002/14651858.CD007459.pub2] [Medline: 23235644]
29. Oh B, Cho B, Han MK, Choi H, Lee MN, Kang H, et al. The effectiveness of mobile phone-based care for weight control in metabolic syndrome patients: randomized controlled trial. JMIR Mhealth Uhealth 2015;3(3):e83 [FREE Full text] [doi: 10.2196/mhealth.4222] [Medline: 26293568]

30. Aranda-Jan CB, Mohutsiwa-Dibe N, Loukanova S. Systematic review on what works, what does not work and why of implementation of mobile health (mHealth) projects in Africa. BMC Public Health 2014;14:188 [FREE Full text] [doi: 10.1186/1471-2458-14-188] [Medline: 24555733]
31. Tran MC, Labrique AB, Mehra S, Ali H, Shaikh S, Mitra M, et al. Analyzing the mobile "digital divide": changing determinants of household phone ownership over time in rural bangladesh. JMIR Mhealth Uhealth 2015;3(1):e24 [FREE Full text] [doi: 10.2196/mhealth.3663] [Medline: 25720457]
32. Rice R, Katz J. Comparing internet and mobile phone usage: digital divides of usage, adoption, and dropouts. Telecommun Policy 2003 Sep;27(8-9):597-623. [doi: 10.1016/S0308-5961(03)00068-5]
33. Abdelfattah BM, Bagchi KK, Udo GJ, Kirs P. Understanding the internet digital divide: an exploratory multi-nation individual-level analysis. : AMCIS; 2010 Jan 08 Presented at: Americas Conference on Information Systems; 2010; Lima, Peru.
34. Chindo L. Mobile phone use in Cameroon: an increasingly useful tool for the follow-up of children with Burkitt lymphoma. Pediatr Blood Cancer 2013 Mar;60(3):524. [doi: 10.1002/pbc.24226] [Medline: 23065945]
35. Parham GP, Mwanahamuntu MH, Pfaendler KS, Sahasrabuddhe VV, Myung D, Mkumba G, Stringer Jeffrey S A. eC3: a modern telecommunications matrix for cervical cancer prevention in Zambia. J Low Genit Tract Dis 2010 Jul;14(3):167-173 [FREE Full text] [doi: 10.1097/LGT.0b013e3181cd6d5e] [Medline: 20592550]
36. Odigie VI, Yusufu LM, Dawotola DA, Ejagwulu F, Abur P, Mai A, et al. The mobile phone as a tool in improving cancer care in Nigeria. Psychooncology 2012 Mar;21(3):332-335. [doi: 10.1002/pon.1894] [Medline: 22383275]
37. Pastakia SD, Karwa R, Kahn CB, Nyabundi JS. The evolution of diabetes care in the rural, resource-constrained setting of western Kenya. Ann Pharmacother 2011 Jun;45(6):721-726. [doi: 10.1345/aph.1P779] [Medline: 21558485]
38. World Health Organization. STEPwise approach to surveillance (STEPS). Geneva, Switzerland: WHO; 2016. URL: http://www.who.int/chp/steps/en/ [accessed 2016-02-02] [WebCite Cache ID 6f03W3gRf]
39. Bovet P, Viswanathan B, Louange M, Gedeon J. Republic of Seychelles National Survey of Noncommunicable Diseases in Seychelles 2013-2014. Victoria, Seychelles: Ministry of Health; 2015.
40. National Bureau of Statistics. Seychelles in Figures. Victoria, Seychelles: National Bureau of Statistics; 2015 Nov. URL: http://www.nsb.gov.sc/wp-content/uploads/2015/12/Seychelles-in-Figures-2015-edition1.pdf [accessed 2016-02-02] [WebCite Cache ID 6f03heVwg]
41. Mardikyan S, Yıldız E, Ordu M, Şimşek B. Examining the global digital divide: a cross-country analysis. CIBIMA 2015 Jan 20:1-10. [doi: 10.5171/2015.592253]
42. Aubert L, Bovet P, Gervasoni JP, Rwebogora A, Waeber B, Paccaud F. Knowledge, attitudes, and practices on hypertension in a country in epidemiological transition. Hypertension 1998 May;31(5):1136-1145 [FREE Full text] [Medline: 9576126]
43. Jennings L, Gagliardi L. Influence of mHealth interventions on gender relations in developing countries: a systematic literature review. Int J Equity Health 2013;12:85 [FREE Full text] [doi: 10.1186/1475-9276-12-85] [Medline: 24131553]
44. Wesolowski A, Eagle N, Noor AM, Snow RW, Buckee CO. Heterogeneous mobile phone ownership and usage patterns in Kenya. PLoS One 2012;7(4):e35319 [FREE Full text] [doi: 10.1371/journal.pone.0035319] [Medline: 22558140]
45. Mo Ibrahim Foundation. Ibrahim Index of African Governance. London, UK: Mo Ibrahim Foundation; 2015. URL: http://mo.ibrahim.foundation/iiag/ [accessed 2016-05-01] [WebCite Cache ID 6hB57sNYC]
46. Loges W, Jung J. Exploring the digital divide: internet connectedness and age. Commun Res 2001 Aug 01;28(4):536-562. [doi: 10.1177/009365001028004007]
47. Brodie M, Flournoy RE, Altman DE, Blendon RJ, Benson JM, Rosenbaum MD. Health information, the Internet, and the digital divide. Health Aff (Millwood) 2000;19(6):255-265 [FREE Full text] [Medline: 11192412]

Abbreviations

NCD: Noncommunicable disease
SES: socioeconomic status
SMS: short message service
WHO: World Health Organization

Edited by G Eysenbach; submitted 08.02.16; peer-reviewed by J Schindler-Ruwisch; comments to author 02.03.16; revised version received 15.03.16; accepted 28.03.16; published 20.05.16.

Please cite as:
Yepes M, Maurer J, Viswanathan B, Gedeon J, Bovet P
Potential Reach of mHealth Versus Traditional Mass Media for Prevention of Chronic Diseases: Evidence From a Nationally Representative Survey in a Middle-Income Country in Africa
J Med Internet Res 2016;18(5):e114
URL: http://www.jmir.org/2016/5/e114/
doi:10.2196/jmir.5592
PMID:27207074

Original Paper

Hearing Tests on Mobile Devices: Evaluation of the Reference Sound Level by Means of Biological Calibration

Marcin Masalski[1,2], MD, PhDEng; Lech Kipiński[3], MD, PhDEng; Tomasz Grysiński[2], PhDEng; Tomasz Kręcicki[1], MD, PhD

[1]Department and Clinic of Otolaryngology, Head and Neck Surgery, Wroclaw Medical University, Wrocław, Poland

[2]Department of Biomedical Engineering, Wroclaw University of Technology, Wrocław, Poland

[3]Department of Pathophysiology, Wroclaw Medical University, Wrocław, Poland

Corresponding Author:
Marcin Masalski, MD, PhDEng
Department and Clinic of Otolaryngology, Head and Neck Surgery
Wroclaw Medical University
Borowska 213
Wrocław, 50-556
Poland
Phone: 48 515086252
Fax: 48 717343709
Email: marcin.masalski@pwr.edu.pl

Abstract

Background: Hearing tests carried out in home setting by means of mobile devices require previous calibration of the reference sound level. Mobile devices with bundled headphones create a possibility of applying the predefined level for a particular model as an alternative to calibrating each device separately.

Objective: The objective of this study was to determine the reference sound level for sets composed of a mobile device and bundled headphones.

Methods: Reference sound levels for Android-based mobile devices were determined using an open access mobile phone app by means of biological calibration, that is, in relation to the normal-hearing threshold. The examinations were conducted in 2 groups: an uncontrolled and a controlled one. In the uncontrolled group, the fully automated self-measurements were carried out in home conditions by 18- to 35-year-old subjects, without prior hearing problems, recruited online. Calibration was conducted as a preliminary step in preparation for further examination. In the controlled group, audiologist-assisted examinations were performed in a sound booth, on normal-hearing subjects verified through pure-tone audiometry, recruited offline from among the workers and patients of the clinic. In both the groups, the reference sound levels were determined on a subject's mobile device using the Bekesy audiometry. The reference sound levels were compared between the groups. Intramodel and intermodel analyses were carried out as well.

Results: In the uncontrolled group, 8988 calibrations were conducted on 8620 different devices representing 2040 models. In the controlled group, 158 calibrations (test and retest) were conducted on 79 devices representing 50 models. Result analysis was performed for 10 most frequently used models in both the groups. The difference in reference sound levels between uncontrolled and controlled groups was 1.50 dB (SD 4.42). The mean SD of the reference sound level determined for devices within the same model was 4.03 dB (95% CI 3.93-4.11). Statistically significant differences were found across models.

Conclusions: Reference sound levels determined in the uncontrolled group are comparable to the values obtained in the controlled group. This validates the use of biological calibration in the uncontrolled group for determining the predefined reference sound level for new devices. Moreover, due to a relatively small deviation of the reference sound level for devices of the same model, it is feasible to conduct hearing screening on devices calibrated with the predefined reference sound level.

(J Med Internet Res 2016;18(5):e130) doi:10.2196/jmir.4987

KEYWORDS

hearing test, mobile device, calibration

Introduction

This paper concerns the calibration of mobile devices for the purposes of the pure-tone audiometry and explores the possibility of the process automation. This study is the first part of the planned research, which aims to define the accuracy of pure-tone audiometry conducted on mobile devices using predefined calibration coefficients.

The evaluation of air conduction hearing threshold may be performed on common electronic devices such as personal computers, laptops, tablets, or mobile phones [1-11]. The examination may not only be applied in screening tests but can also be useful for self-monitoring in the following hearing disorders: fluctuating hearing loss, Menier's disease, tinnitus, sudden sensorineural hearing loss, age-related hearing loss, or during ototoxic therapy [1,2]. Moreover, it may prove beneficial for preliminary evaluation of patients with otologic complaints [3], by saving the time needed for more sophisticated examinations or for patients who require more medical attention [12]. Mobile phone–based hearing tests can also complement other telehealth methods in otolaryngology such as remote hearing aid adjustment [13] or mobile phone–based otoendoscopy [14]. Hearing examinations conducted on common electronic devices may turn out to be particularly useful in these parts of the world where access to hearing health care professionals is severely limited [15,16].

The air conduction hearing threshold of the pure-tone audiometry can be compared with the self-determined threshold carried out in home setting on common electronic devices in terms of the measurement algorithm, environmental noise, and the calibration method. The accuracy of the automated algorithms for the assessment of the hearing threshold is comparable to that of the ascending method of the pure-tone audiometry [3,5,10,17-19]. Examinations carried out in quiet rooms are characterized by a similar error as those carried out in sound booths [3,4], particularly in patients with hearing impairments [20].

However, the accuracy of hearing examinations conducted in home setting is highly dependent on the calibration method. The device calibration for home-applied self-assessed hearing test may be conducted in various ways, but its omission leads to significant measurement errors [6-9,21]. The device and its components may be calibrated in a laboratory [3-5,10,22,23]. Alternatively, calibration may be conducted in home settings by means of the biological method, that is, in relation to the hearing threshold of a normal-hearing person [1,2,11,24]. The devices calibrated in laboratory conditions exhibit the smallest error, and for those devices, there is no significant statistical difference compared with the pure-tone audiometry [4,5,10,22]. Laboratory calibration may also be conducted for groups of devices that are standardized with similar hardware and software components. In the case of Apple iOS–based devices, the differences in sound intensity between different sets are within 4 dB, making it possible to achieve accurate results in comparison with conventional audiometry [3]. By far, the biggest error is found in the case of biological calibration. The mean standard error of the biological calibration method based on the Bekesy audiometry is estimated at 4.90 dB [24]. However, biological calibration does not require specialized equipment, and its application significantly improves the accessibility of the examination.

Popularization of mobile devices that are offered with bundled headphones entails the possibility of using the once-determined reference sound level in all sets of the same type. Due to the rapidly growing market of mobile devices and their diversity, it seems appropriate to propose a method of automatic determination of the reference sound level without having to calibrate each new device model in a laboratory. The predefined reference sound level determined by means of the biological calibration can meet the aforementioned requirements. Until reliable reference level has been obtained, a user will be asked to perform the biological calibration before the test. This method is particularly important for Android-based devices, which account for about 80.7% of the mobile devices market (data for 2014) [25]. Contrary to iOS-based devices, they do not form a homogeneous group owing to the wide variety of hardware solutions. It can be assumed that tests carried out on the basis of the predefined reference sound level will be more accurate than those based on a single biological calibration. At the same time, the scalability of the proposed solution concerning the number of device models is greater than a laboratory calibration, and therefore, the availability of the test is expected to be higher as well.

In this paper, the reference sound level was compared between the uncontrolled and controlled groups. In the uncontrolled group, the measurements were conducted at home by the users themselves to prepare for further hearing tests. In the controlled group, the audiologist-assisted calibration was performed in a sound booth, in a group of normal-hearing persons verified through the pure-tone audiometry. The test results were compared between the groups to examine the possibility of using the uncontrolled measurements to determine predefined reference sound levels.

Methods

It was a single-center, parallel trial with no randomization carried out on 2 separate groups of participants: controlled and uncontrolled.

In the uncontrolled group, subjects were recruited via an open access, free mobile phone app "Hearing Test" [26] available on Google Play, designed for conducting hearing examinations on mobile devices. The functionality of the app was presented on the Google Play website so that the participants could learn more before the installation. A prerequisite for using the app was giving consent during the initial launch for the data to be anonymously used for the purposes of app development and scientific work. The main functionality of the app is the assessment of the hearing threshold within the frequency band 250 Hz to 8 kHz in relation to the reference sound level. If the predefined reference sound level for the particular model was not available in the centralized database, the device had to be calibrated before the examination. Calibration could also be conducted to verify the reference sound level used during the examination. Therefore, the calibration measurements, which

are subject of interest in this paper, were not the outcome of intended actions but resulted from the wish to conduct the hearing test or to verify the test results. All the users who conducted calibration using the bundled headphones were assigned to the uncontrolled group. Detailed instructions on calibration were presented directly before the measurement. They contained the requirements of conducting the test in silence, on headphones, by a normal-hearing person aged 18 to 35 years. Subjects were identified by means of a unique number assigned to the device, more precisely to the instance of the device's operational system. When more than 1 calibration was assigned to the same identification number, the results of the last calibration were analyzed. The device model was identified by the manufacturer name and the end-user-visible name for the end product.

In the controlled group, the participants were recruited offline from among the employees and patients of otolaryngology clinic through face-to-face prompting. The eligibility criteria were the possession of an Android-based mobile phone with bundled headphones and willingness to participate in the study. In the first stage, the hearing threshold of the participant was verified through the pure-tone audiometry with the ascending method [27] conducted on a clinical audiometer Interacoustic AD229e with TDH-39 headphones previously calibrated according to the norm ISO 389-1:1998. Having installed the "Hearing Test" app, the calibration was performed in a sound booth with the use of bundled headphones. If the hearing of the participant was within the normal range (ie, no higher than 20 dB HL in the frequency between 125 Hz and 8 kHz), the participant was asked to perform calibration twice under the supervision of an audiologist. If not, calibration was conducted twice by 1 of 2 normal-hearing audiologists. The results of the first calibration were used in the analysis, whereas the other one served for calculating the test–retest difference.

In both the controlled and uncontrolled groups, calibration consisted of determining the hearing threshold by means of the Bekesy audiometry. The Bekesy audiometry is a self-recording hearing test in which the subject controls the intensity of the stimulus by pressing a button, while the frequency of the stimulus is being slowly changed within the audible range. The intensity of a stimulus decreases as long as the button is pressed and increases when released, yielding a near-threshold zigzag tracing. The calibration was carried out using a pulsed tone, within the frequency band 125 Hz to 16 kHz, at frequency change 1 octave/minute and intensity change 2 dB/second [24]. It took 7 minutes to complete the calibration. The results were stored in the centralized database. Calibration coefficients were determined for 250 Hz, 500 Hz, 1 kHz, 2 kHz, 4 kHz, 6 kHz, and 8 kHz by calculating the median within ±1/2 octave of sound intensities, at which a change in the button state occurred. Thus, the calibration coefficients determine the level of the signal generated by the device, which equals the hearing threshold of the reference person.

Comparison of Reference Sound Levels

The reference sound level has been defined as such a level of the signal generated by the device, which will produce in the bundled headphones the sound at the intensity of 0 dB HL. The reference sound level can be estimated based on calibration coefficients assuming that the hearing threshold of the reference person nears 0 dB HL. The estimation will be much more accurate if the calibration coefficients are decreased by the hearing threshold of the reference person. By the same token, the reference sound level can be estimated on the basis of population if the distribution of its calibration coefficients and the hearing threshold are known.

The groups were compared by means of the reference sound level. In the uncontrolled group, the reference sound level was obtained by decreasing calibration coefficients by the literature-based median of the hearing threshold in the population screened for ear-related disorders and a history of noise exposure [28]. In the controlled group, the reference sound level was calculated as the mean value of calibration coefficients decreased by the pure-tone hearing thresholds. The comparison of reference sound levels in groups was conducted on the basis of confidence intervals (CIs).

The sample size was determined on the basis of the standard deviation (SD) of the difference between reference sound levels, and therefore, it constitutes a set of alternative pairs of numbers determining the required number of calibrations in the controlled and uncontrolled groups. The more calibrations are conducted in one group, the fewer will be required in the other. On the basis of preliminary measurements, for statistical significance .05, statistical power 0.8, and the magnitude of the effect 15 dB, the required number of calibrations for a single model in the uncontrolled group was determined at 60, 25, or 15 on the assumption of 1, 2, or 3 calibrations in the controlled group, respectively. The experiment was completed after satisfying criteria for 10 different models. The number of models was agreed arbitrarily.

Intramodel and Intermodel Analyses

Additional analyses, that are, intramodel and intermodel analyses, have been carried out. The accuracy of the hearing examination conducted on the basis of predefined reference sound level is dependent on the variability of reference sound level in the group of devices belonging to the same model. On the basis of the results in the controlled group, an attempt was made to estimate the aforementioned variability.

Moreover, the necessity of applying different sound reference levels for different models has been verified. The comparison of the reference sound levels between models was performed based on the data in the uncontrolled group.

The "Hearing Test" app was developed based on Web page–embedded Java applet ported to Android system. The Java applets had previously been used in the research on validity evaluation of self-test Web-based pure-tone audiometry [2] and in the research on biological calibration methods for Web-based hearing tests [24]. During the trial, based on preliminary analyses, the number of calibrations required to calculate predefined calibration coefficients was reduced (from 100 to 16; September 6, 2014). Although the change caused a slower growth of the number of calibrations in the uncontrolled group for models that have exceeded the required threshold, it contributed to improving the app's ratings in Google Play,

increasing the number of users, and thus quickly obtaining the required sample size. The remaining modifications to the app were connected with its development (September 12, 2014) and correction of minor bugs (September 25, 2014). The app's functionalities were expanded to include the possibility of adding notes, sending and printing the examination results, and adjusting the calibration coefficients on the basis of the pure-tone audiogram. During the trial, the app was made available in 8 new languages besides English and Polish also in German, Spanish, French, Italian, Japanese, Korean, Portuguese, and Russian (September 12, 2014).

Results

In the period between November 30, 2013 and February 13, 2015, the app was installed 114,546 times from Google Play. During that time, 20,747 calibrations were carried out on 18,154 different devices representing 2772 different models. Overall, 8988 of these calibrations carried out on 8620 different devices (2044 different models) were performed on the bundled headphones. In the controlled group, 158 calibrations (test and retest) were conducted on 79 devices representing 50 models. All calibrations conducted on the remaining 8541 devices representing 2040 models, that is, 8830 calibrations, were qualified to the uncontrolled group. In the controlled group, 11 of 79 participants (14%) were found to have slight hearing impairment, and in these cases, the calibration was performed by 1 of the 2 normal-hearing audiologists. The analysis was performed on 10 models for which the required sample size was obtained, that is, at least 60, 25, or 15 calibrations in the uncontrolled group in the case of 1, 2, or 3 calibrations in the controlled group, respectively (Figure 1).

Figure 1. Flow diagram.

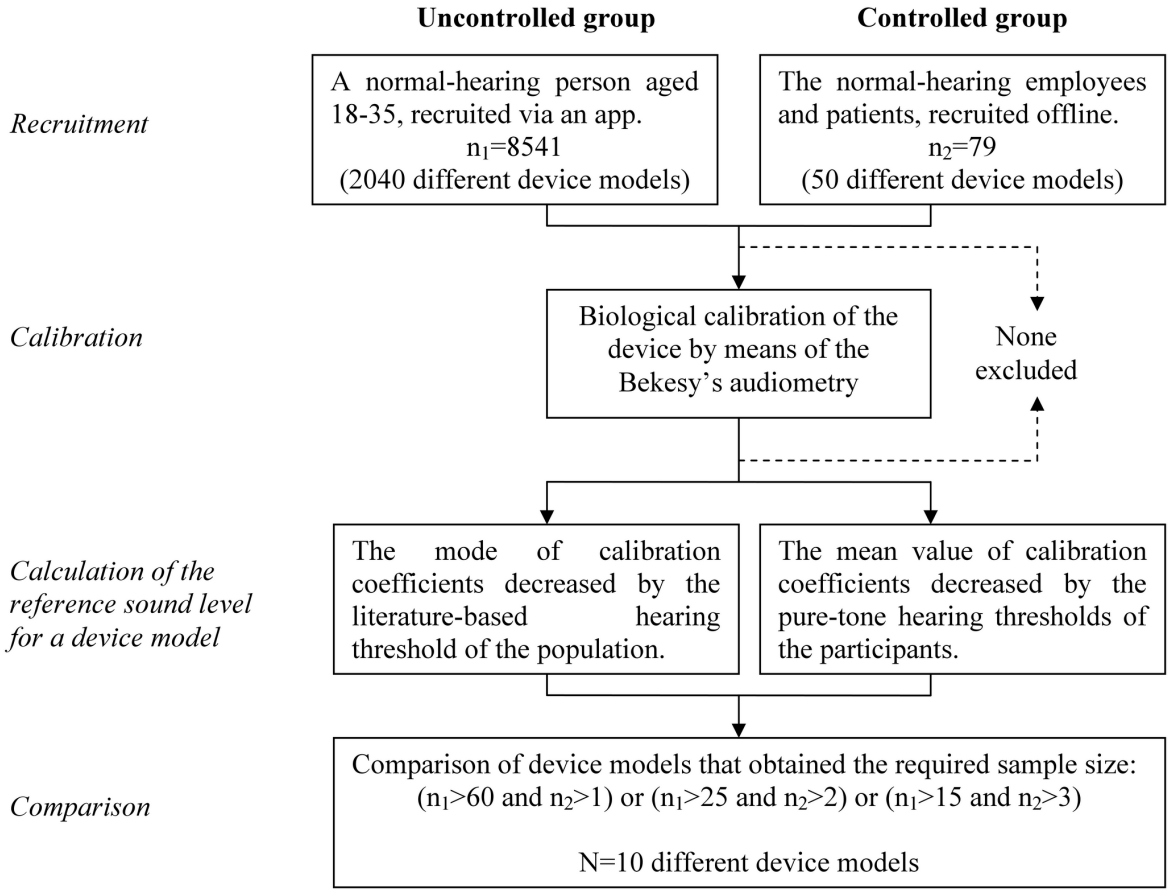


Figure 1. Flow diagram.

Comparison of Reference Sound Levels

In the uncontrolled group, the central tendency of calibration coefficients was estimated by the mode as it is more robust than the median to outliers [29-31]. Hence, no data were excluded from further analysis. Choosing the mode as a measure of the central tendency was dictated by significant asymmetry in the distribution (Figure 2), whose causes should be sought in greater probability of error, leading to overestimation of the coefficient rather than its underestimation (see Discussion). The mode was determined by the mean of shortest half sample [29] and its SD by the bootstrap method [32].

In the uncontrolled group, the reference sound level was obtained by decreasing the mode of calibration coefficient by

the median of the hearing threshold level for the population meeting the calibration requirements, that is, aged 18 to 35 years with no prior hearing problems. The median was estimated on the literature-based data of the hearing threshold in population screened for ear-related disorders and a history of noise exposure [28]. The median for the group aged 18 to 35 years was calculated as the weighted arithmetic mean under the assumption that the hearing threshold in the group aged 18 to 19 years is similar to that in the group aged 20 to 29 years (Table 1). Due to considerable size of the trial sample (n=5498) used to determine the median value, its error was considered negligible.

In the controlled group, the reference sound levels were determined by decreasing the values of calibration coefficients by the pure-tone hearing thresholds. The SD of the reference sound levels was estimated by means of the differences between devices of the same model. The differences were calculated at all 7 frequencies. The SD of the difference was obtained at the level of 9.09 dB (95% CI 8.67-9.55), and thus, the SD of the reference sound level in the controlled group was estimated at 6.43 dB (95% CI 6.13-6.75).

The reference sound levels in both the groups are presented in Figure 3. They are expressed by the intensity of digital signals, which generate sounds at the level of 0 dB HL. It has been arbitrarily assumed that the lowest possible signal level that can be theoretically generated, whose intensity results directly from quantization, will be at the level of −40 dB. With the aforementioned assumptions, for most mobile devices, the signals with the frequency of 1 kHz and intensity of 0 dB generate sounds close to the level of 0 dB HL.

Reference sound levels in the uncontrolled group were compared with the reference sound levels in the controlled group (Figure 3,Tables 2 and 3). CIs were estimated on the basis of the SD of the difference between these values (Multimedia Appendix 1). At 95% CIs, statistically significant differences were found in 5 of 77 (7%), which is in accordance with the statistical significance level and confirms the similarity of reference sound levels in both the groups. The mean value of the differences calculated jointly for all the frequencies and all the analyzed models was 1.50 dB (SD 4.42).

Table 1. Literature-based medians of the hearing threshold by gender and groups aged 20 to 29 years and 30 to 39 years [28] and the estimated median of the hearing threshold for the group aged 18 to 35 years.

			Hearing threshold median (dB HL)						
Gender, age	N	Weight	250 Hz	500 Hz	1 kHz	2 kHz	4 kHz	6 kHz	8 kHz
Men, 20-29	650	12	12.5	5.0	2.5	2.5	5.0	10.0	7.5
Women, 20-29	1840	12	12.5	7.5	2.5	2.5	2.5	10.0	7.5
Men, 30-39	619	5	12.5	7.5	5.0	5.0	10.0	15.0	10.0
Women, 30-39	2389	5	15.0	7.5	5.0	5.0	5.0	12.5	10.0
Estimated median, 18-35.			12.9	6.6	3.2	3.2	4.9	11.1	8.2

Table 2. Differences in reference sound levels between uncontrolled and controlled groups by device models.

Model	Number of calibrations in the uncontrolled group	Number of calibrations in the controlled group	Difference between reference sound levels in groups, dB (SD)
HTC ONE	99	2	−0.75 (4.58)
SAMSUNG GT-I8190	43	5	4.99 (5.16)
SAMSUNG GT-I9100	94	1	1.29 (3.92)
SAMSUNG GT-I9105P	26	3	3.35 (3.06)
SAMSUNG GT-I9195	83	3	5.80 (4.25)
SAMSUNG GT-I9300	100	7	−0.04 (4.29)
SAMSUNG GT-I9505	108	5	0.45 (4.43)
SAMSUNG GT-N7100	88	1	4.78 (4.28)
SONY C6603	62	2	−1.14 (4.31)
SONY C6903	55	2	−3.73 (5.94)

Table 3. Differences in reference sound levels between uncontrolled and controlled groups by frequencies (number of models n=10).

Frequency (Hz)	Difference between reference sound levels in groups, dB (SD)
250	1.58 (5.91)
500	2.63 (4.59)
1k	4.68 (2.86)
2k	2.15 (4.41)
4k	0.53 (4.13)
6k	0.67 (6.48)
8k	−1.74 (5.72)

Figure 2. Calibration coefficients for sample device models at 1 kHz in the uncontrolled group.

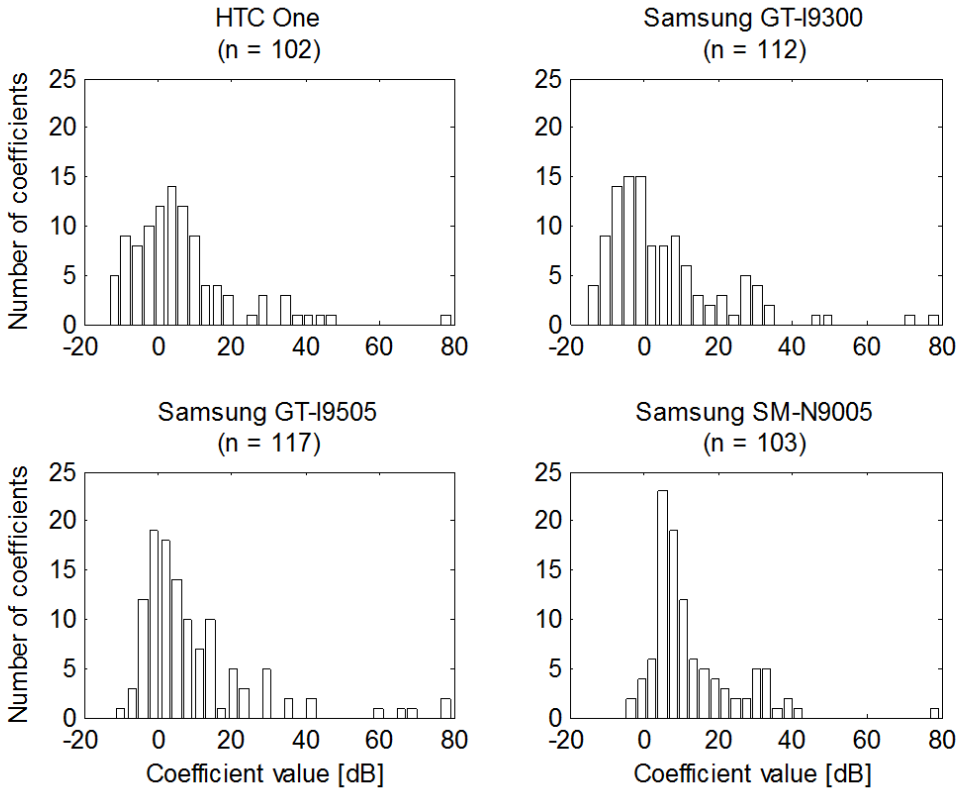
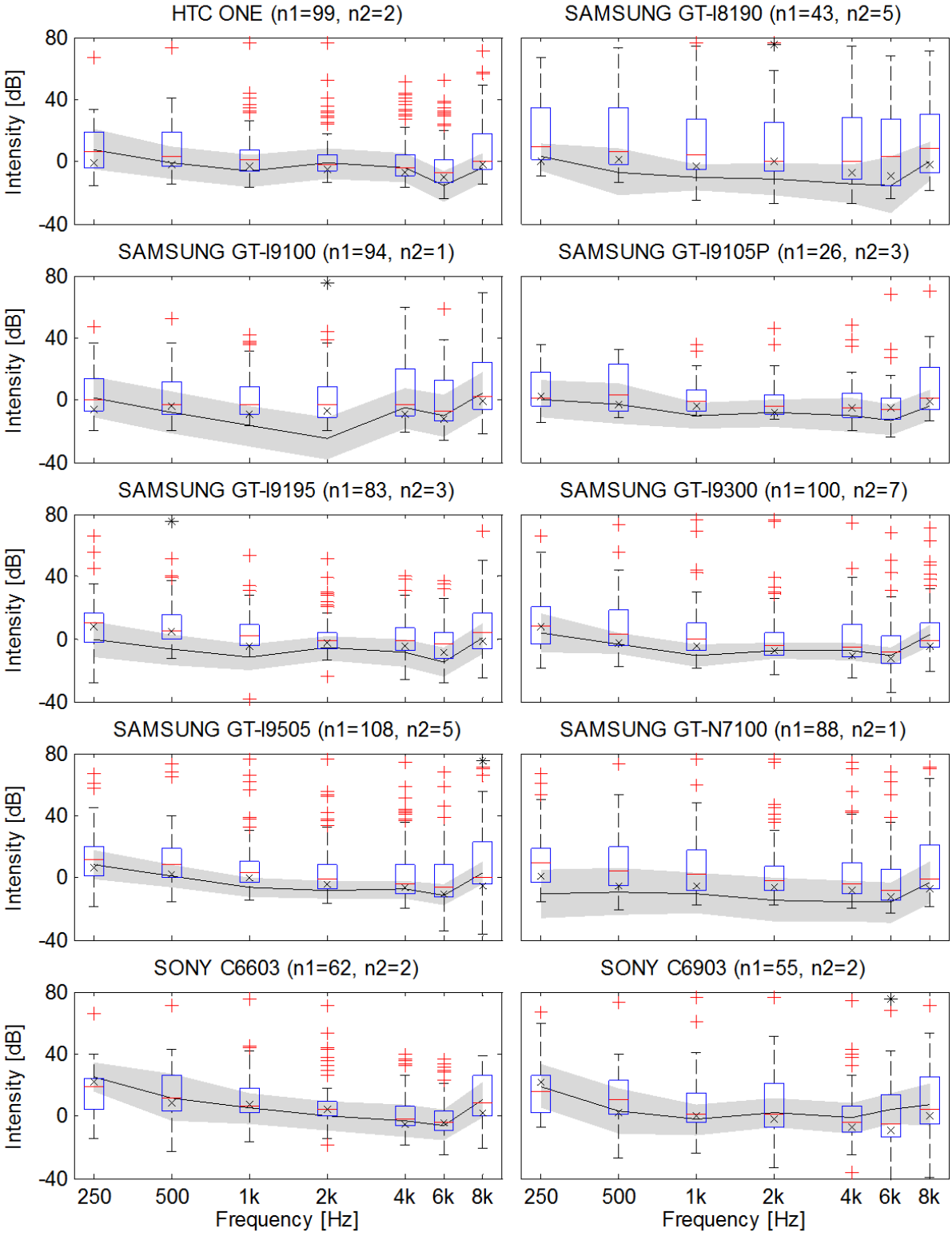


Figure 3. The reference sound level for different device models (boxplot—calibration coefficients in the uncontrolled group decreased by the median of the hearing threshold in the population of normal-hearing subjects, x—the reference sound level in the uncontrolled group, that is, the mode of boxplot data, continuous line—the reference sound level in the controlled group, gray area—95% confidence interval for the difference between reference sound level in groups, *—statistically significant difference).



Intramodel Analysis

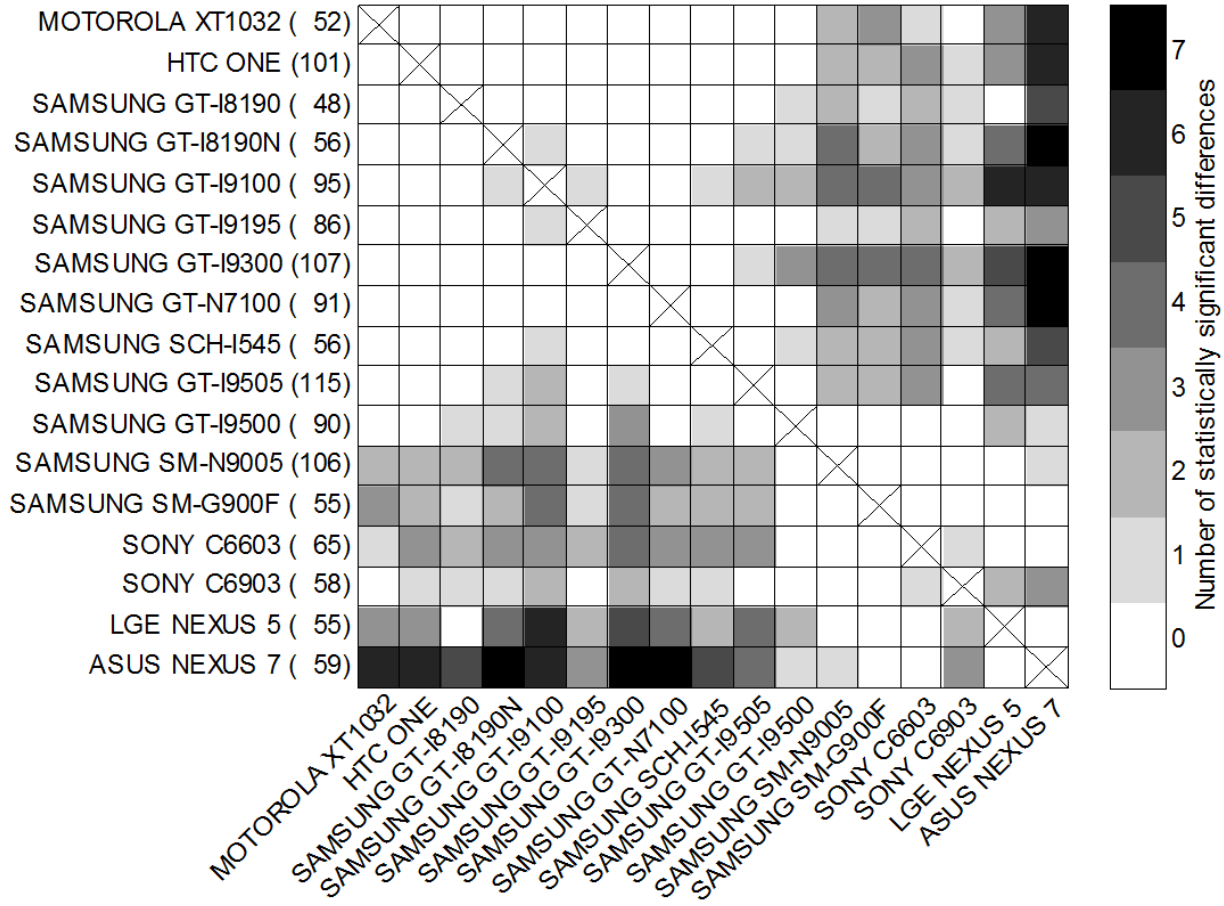
The reference sound level in the controlled group was measured by decreasing the value of calibration coefficient by the pure-tone hearing threshold of the reference person. Therefore, the variability of the measured value of the reference sound level is dependent on the variability of the real value of the reference sound level, the measurement error of the pure-tone

hearing threshold, and the measurement error of calibration coefficient. Assuming the literature-based SD of the test–retest examination for pure-tone audiometry at 5.37 dB (95% CI 5.02-5.77) [24], the error of determining pure-tone hearing threshold is 3.80 dB (95% CI 3.55-4.08). The measurement error of the calibration coefficient was determined directly on the basis of the obtained results. The SD in the test–retest examination was found to be 4.62 dB (95% CI 4.37-4.91), and thus, the error is 3.27 dB (95% CI 3.09-3.48). Finally, using the variability of the measured value of the reference sound level described in the previous paragraph, the variability of the real value of the reference sound level was expressed by SD of 4.03 dB (95% CI 3.93-4.11; Multimedia Appendix 2).

Intermodel Analysis

Reference sound levels were compared between models based on the measurements carried out in the uncontrolled group. The distributions of differences between modes of reference sound levels were determined using the bootstrap method. The comparison was conducted at the level of statistical significance P=.01 based on the critical value determined by SD of the bootstrap distribution. Statistically significant differences were found in reference sound levels (Figure 4).

Figure 4. Comparison of reference sound levels between models by means of the number of frequencies with statistically significant differences (P=.01).



Discussion

Reference sound levels determined for the purpose of the hearing test were compared between controlled and uncontrolled groups for different models of mobile devices. The difference between reference sound levels was obtained at 1.50 dB (SD 4.42), which confirms the possibility of using calibration in the uncontrolled group for determining predefined reference sound level for new devices.

The mean, real variability of the reference sound level in the group of devices belonging to the same model was expressed by the SD estimated at 4.03 dB (95% CI 3.93-4.11). This value limits the hearing test accuracy when a common, predefined reference sound level is applied to all devices of the same model. The estimated variability of reference sound level is higher than the value determined for iPads, for which the maximum intrafrequency difference was 4 dB [3]. One of the reasons for such a difference could be the change in frequency characteristics of headphones during their usage. In the conducted measurements, new headphones were used as well as long-used ones. The variability determined previously is, however, significantly smaller than the error of the biological method based on the Bekesy audiometry estimated at the level of 4.90 dB (95% CI 4.46-5.45) [24].

Statistically significant differences were found in the reference sound levels between models, in particular between models produced by different manufacturers. At the same time, some models had similar reference sound levels. This confirmed the

need to determine reference levels individually for various groups of devices, which does not necessarily correspond with the proposed division that was based on manufacturer of the product and the end-user-visible name for the end product.

Distribution of calibration coefficients in the uncontrolled group was characterized by significant asymmetry caused by higher number of coefficients with potentially overstated values. This may be related to greater odds of making a mistake leading to overstating coefficient values rather than understating them. The overstated value of the coefficient may be caused by potential errors such as noise during calibration, hearing deficits of the reference person, calibration on damaged headphones, or abandoning the device during calibration. The understated coefficient values may only be influenced by interferences or other disruptions when they are misinterpreted as a reference signal. The mean difference of the reference sound level between groups calculated on the basis of the mode was 1.50 dB (SD 4.42). However, if a median is used instead of a mode, the mean difference will be 5.39 dB (SD 3.93), which proves the validity of the applied measure of the central tendency.

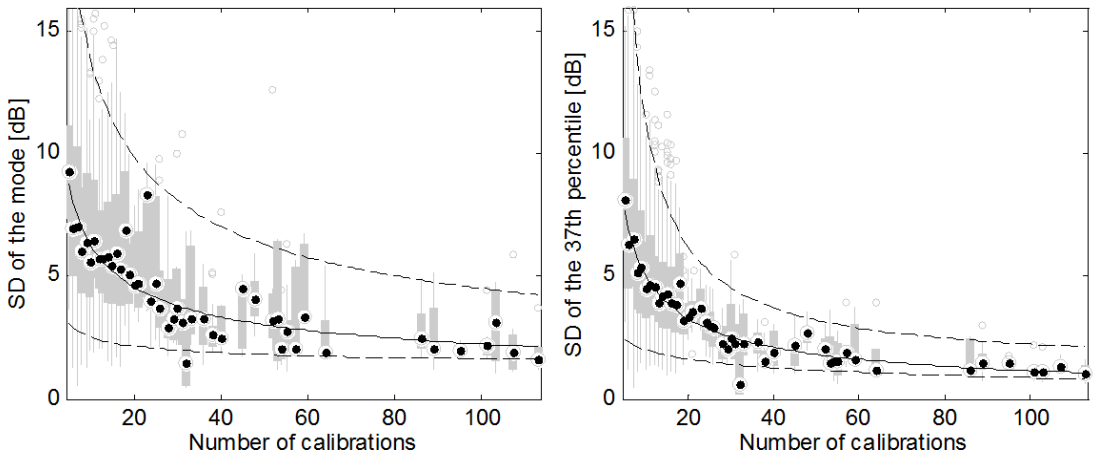
Predefined reference sound levels for new device models may be determined semiautomatically based on the calibrations conducted in the uncontrolled group. To estimate the error of the determination, the SD of the mode of reference sound level was calculated in relation to the number of calibrations (Figure 5). Calculations were carried out for every model, cumulatively for all frequencies, using the bootstrap method. The mean standard error of determining the predefined reference sound level was found below 5 dB for the number of calibrations greater than 16.

The mode value precisely determines the reference sound level but turns out to be not stable enough with a small number of calibrations. Therefore, an attempt has been made to determine the reference sound level using a quantile, which in most cases included the mode. The modal quantile has been found at the 37th percentile. The mean difference between the groups, determined on the basis of the 37th percentile, is comparable to the value obtained for the mode and equals 1.02 dB (SD 3.99). At the same time, the 37th percentile is characterized by significantly smaller SD (Figure 5). This is particularly important in the case of a small number of calibrations. For the 37th percentile, the mean standard error of determining the predefined reference sound level was found below 5 dB for the number of calibrations greater than 10.

The accuracy of hearing examinations determines their application. Therefore, an attempt was made at estimating the SD of the difference between classical pure-tone audiometry and a hearing test conducted on a mobile device calibrated by means of predefined sound level that was determined in uncontrolled settings using the biological method. Assuming the standard error of determining the predefined reference sound level for 16 calibrations at 5.08 dB (95% CI 2.25-11.02), variability of reference sound level within a single model at 4.03 dB (95% CI 3.93-4.11), the literature-based SD of test–retest differences of the pure-tone audiometry at 5.37 dB (95% CI 5.02-5.77) [24] and making a conservative assumption of independence of the aforementioned variables, we obtain the SD of the difference at 8.42 dB (95% CI 6.76-13.16) (Multimedia Appendix 3). The aforementioned value is comparable to the values of the screening methods, which had been presented in earlier works [1,2] and thereby validates the application of the method in hearing screening.

The method presented in this paper can be applied in screening hearing examinations on a large scale with the use of popular mobile devices sold with bundled headphones. Due to rapidly growing market of mobile devices, the main advantage of the method is the semiautomated calibration of new models. Predefined reference sound level for a new model may be determined on the basis of a biological calibration conducted by the first users of devices. To confirm the estimated accuracy of the method, it is advisable to conduct a direct comparison of pure-tone audiometry and a hearing test on mobile devices calibrated biologically by means of the predefined reference sound level.

Figure 5. Standard deviation of reference sound level in relation to the number of calibrations (continuous line—median; dotted lines—percentiles: 2.5% and 97.5%).



Acknowledgments

The study was financed by Wroclaw Medical University as part of Research Grant for Young Scientists (grant no. 170/Pbmn).

Conflicts of Interest

The first author of this paper is the author and owner of the “Hearing Test” app, by means of which the measurements described in this paper were conducted.

Multimedia Appendix 1

Standard Deviation of the Difference Between the Mode of the Reference Sound Level in the Uncontrolled Group and the Mean Reference Sound Level in the Controlled Group.

[PNG File, 65KB - jmir_v18i5e130_app1.png]

Multimedia Appendix 2

The Real Variability of the Reference Sound Level Among Devices of the Same Model.

[PNG File, 51KB - jmir_v18i5e130_app2.png]

Multimedia Appendix 3

Estimation of the Standard Deviation of the Difference Between the Pure-Tone Audiometry and a Hearing Test Conducted on a Mobile Device Calibrated by Means of the Predefined Reference Sound Level.

[PNG File, 30KB - jmir_v18i5e130_app3.png]

References

1. Honeth L, Bexelius C, Eriksson M, Sandin S, Litton J, Rosenhall U, et al. An internet-based hearing test for simple audiometry in nonclinical settings: preliminary validation and proof of principle. Otol Neurotol 2010 Jul;31(5):708-714. [doi: 10.1097/MAO.0b013e3181de467a] [Medline: 20458255]
2. Masalski M, Kręcicki T. Self-test web-based pure-tone audiometry: validity evaluation and measurement error analysis. J Med Internet Res 2013;15(4):e71 [FREE Full text] [doi: 10.2196/jmir.2222] [Medline: 23583917]
3. Foulad A, Bui P, Djalilian H. Automated audiometry using apple iOS-based application technology. Otolaryngol Head Neck Surg 2013 Nov;149(5):700-706. [doi: 10.1177/0194599813501461] [Medline: 23963611]
4. Kam Anna Chi Shan, Sung John Ka Keung, Lee T, Wong Terence Ka Cheong, van HA. Clinical evaluation of a computerized self-administered hearing test. Int J Audiol 2012 Aug;51(8):606-610. [doi: 10.3109/14992027.2012.688144] [Medline: 22676699]
5. Swanepoel DW, Myburgh HC, Howe DM, Mahomed F, Eikelboom RH. Smartphone hearing screening with integrated quality control and data management. Int J Audiol 2014 Dec;53(12):841-849. [doi: 10.3109/14992027.2014.920965] [Medline: 24998412]
6. Szudek J, Ostevik A, Dziegielewski P, Robinson-Anagor J, Gomaa N, Hodgetts B, et al. Can Uhear me now? Validation of an iPod-based hearing loss screening test. J Otolaryngol Head Neck Surg 2012 Apr;41 Suppl 1:S78-S84. [Medline: 22569055]
7. Khoza-Shangase K, Kassner L. Automated screening audiometry in the digital age: exploring uhear™ and its use in a resource-stricken developing country. Int J Technol Assess Health Care 2013 Jan;29(1):42-47. [doi: 10.1017/S0266462312000761] [Medline: 23298579]
8. Handzel O, Ben-Ari O, Damian D, Priel MM, Cohen J, Himmelfarb M. Smartphone-based hearing test as an aid in the initial evaluation of unilateral sudden sensorineural hearing loss. Audiol Neurootol 2013;18(4):201-207. [doi: 10.1159/000349913] [Medline: 23689282]
9. Abu-Ghanem S, Handzel O, Ness L, Ben-Artzi-Blima M, Fait-Ghelbendorf K, Himmelfarb M. Smartphone-based audiometric test for screening hearing loss in the elderly. Eur Arch Otorhinolaryngol 2016 Feb;273(2):333-339. [doi: 10.1007/s00405-015-3533-9] [Medline: 25655259]
10. Van Tasell Dianne J, Folkeard P. Reliability and accuracy of a method of adjustment for self-measurement of auditory thresholds. Otol Neurotol 2013 Jan;34(1):9-15. [doi: 10.1097/MAO.0b013e318278c05d] [Medline: 23202154]
11. McPherson B, Law M M S, Wong M S M. Hearing screening for school children: comparison of low-cost, computer-based and conventional audiometry. Child Care Health Dev 2010 May;36(3):323-331. [doi: 10.1111/j.1365-2214.2010.01079.x] [Medline: 20507326]
12. Margolis RH, Morgan DE. Automated pure-tone audiometry: an analysis of capacity, need, and benefit. Am J Audiol 2008 Dec;17(2):109-113. [doi: 10.1044/1059-0889(2008/07-0047)] [Medline: 18840703]

13. Penteado SP, Bento RF, Battistella LR, Silva SM, Sooful P. Use of the satisfaction with amplification in daily life questionnaire to assess patient satisfaction following remote hearing aid adjustments (telefitting). JMIR Med Inform 2014;2(2):e18 [FREE Full text] [doi: 10.2196/medinform.2769] [Medline: 25599909]
14. Wu C, Wu S, Chen P, Lin Y. An innovative smartphone-based otorhinoendoscope and its application in mobile health and teleotolaryngology. J Med Internet Res 2014;16(3):e71 [FREE Full text] [doi: 10.2196/jmir.2959] [Medline: 24590187]
15. Swanepoel DW, Clark JL, Koekemoer D, Hall JW, Krumm M, Ferrari DV, et al. Telehealth in audiology: the need and potential to reach underserved communities. Int J Audiol 2010 Mar;49(3):195-202. [doi: 10.3109/14992020903470783] [Medline: 20151929]
16. Fagan JJ, Jacobs M. Survey of ENT services in Africa: need for a comprehensive intervention. Glob Health Action 2009;2:n/a [FREE Full text] [doi: 10.3402/gha.v2i0.1932] [Medline: 20027268]
17. Swanepoel DW, Mngemane S, Molemong S, Mkwanazi H, Tutshini S. Hearing assessment-reliability, accuracy, and efficiency of automated audiometry. Telemed J E Health 2010 Jun;16(5):557-563. [doi: 10.1089/tmj.2009.0143] [Medline: 20575723]
18. Ho Allan Thiam Poh, Hildreth AJ, Lindsey L. Computer-assisted audiometry versus manual audiometry. Otol Neurotol 2009 Oct;30(7):876-883. [doi: 10.1097/MAO.0b013e3181b120d0] [Medline: 20179426]
19. Margolis RH, Glasberg BR, Creeke S, Moore Brian C J. AMTAS: automated method for testing auditory sensitivity: validation studies. Int J Audiol 2010 Mar;49(3):185-194. [doi: 10.3109/14992020903092608] [Medline: 20109081]
20. Wong TW, Yu TS, Chen WQ, Chiu YL, Wong CN, Wong A H S. Agreement between hearing thresholds measured in non-soundproof work environments and a soundproof booth. Occup Environ Med 2003 Sep;60(9):667-671 [FREE Full text] [Medline: 12937188]
21. Kimball SH. Inquiry into online hearing test raises doubts about its validity. The Hearing Journal 2008;61(3):38-46. [doi: 10.1097/01.HJ.0000314717.42321.4a]
22. Choi JM, Lee HB, Park CS, Oh SH, Park KS. PC-based tele-audiometry. Telemed J E Health 2007 Oct;13(5):501-508. [doi: 10.1089/tmj.2007.0085] [Medline: 17999612]
23. Senses Examination Platform. URL: http://platformabadanzmyslow.pl/en/about.html [accessed 2012-09-24] [WebCite Cache ID 6AvYBYhVy]
24. Masalski M, Grysiński T, Kręcicki T. Biological calibration for web-based hearing tests: evaluation of the methods. J Med Internet Res 2014;16(1):e11 [FREE Full text] [doi: 10.2196/jmir.2798] [Medline: 24429353]
25. Gartner Newsroom.: Gartner, Inc; 2015 Mar 03. Gartner Says Smartphone Sales Surpassed One Billion Units in 2014 URL: http://www.gartner.com/newsroom/id/2996817, [accessed 2016-05-09] [WebCite Cache ID 6ZyNWSMvX]
26. Masalski M. Hearing Test App - Android Apps on Google Play.: Google Play; 2014. URL: https://play.google.com/store/apps/details?id=mobile.eaudiologia&hl=en [accessed 2016-05-09] [WebCite Cache ID 6TZUfFaE9]
27. British Society of Audiology. 2011. Pure-tone air-conduction and bone-conduction threshold audiometry with and without masking URL: http://www.thebsa.org.uk/resources/pure-tone-air-bone-conduction-threshold-audiometry-without-masking/ [accessed 2016-05-09] [WebCite Cache ID 6hMTzBR56]
28. Engdahl B, Tambs K, Borchgrevink HM, Hoffman HJ. Screened and unscreened hearing threshold levels for the adult population: results from the Nord-Trøndelag Hearing Loss Study. Int J Audiol 2005 Apr;44(4):213-230. [Medline: 16011050]
29. Rousseeuw P, Leroy A. A robust scale estimator based on the shortest half. Statistica Neerland 1988 Jun;42(2):103-116. [doi: 10.1111/j.1467-9574.1988.tb01224.x]
30. Bickel DR. Robust estimators of the mode and skewness of continuous data. Computational Statistics & Data Analysis 2002 Apr;39(2):153-163. [doi: 10.1016/S0167-9473(01)00057-3]
31. Bickel DR. Robust and efficient estimation of the mode of continuous data: the mode as a viable measure of central tendency. Journal of Statistical Computation and Simulation 2003 Dec;73(12):899-912. [doi: 10.1080/0094965031000097809]
32. Efron B, Tibshirani R. An introduction to the bootstrap. New York: Chapman & Hall; 1993.

Abbreviations

CI: confidence interval
HL: hearing level
SD: standard deviation

Edited by G Eysenbach; submitted 28.09.15; peer-reviewed by A Chi Shan Kam, P Moo Kyun; comments to author 29.10.15; revised version received 25.11.15; accepted 21.03.16; published 30.05.16.

Please cite as:
Masalski M, Kipiński L, Grysiński T, Kręcicki T
Hearing Tests on Mobile Devices: Evaluation of the Reference Sound Level by Means of Biological Calibration
J Med Internet Res 2016;18(5):e130
URL: http://www.jmir.org/2016/5/e130/
doi:10.2196/jmir.4987
PMID:27241793

Original Paper

Exploring the Utility of Web-Based Social Media Advertising to Recruit Adult Heavy-Drinking Smokers for Treatment

Krysten W Bold[1], PhD; Tess H Hanrahan[1], M.Res.; Stephanie S O'Malley[1,2], PhD; Lisa M Fucito[1,2,3], PhD

[1]Yale School of Medicine, Department of Psychiatry, New Haven, CT, United States

[2]Yale Cancer Center, New Haven, CT, United States

[3]Smilow Cancer Hospital at Yale-New Haven, New Haven, CT, United States

Corresponding Author:
Krysten W Bold, PhD
Yale School of Medicine
Department of Psychiatry
Connecticut Mental Health Center
34 Park Street
New Haven, CT, 06519
United States
Phone: 1 2039747603
Fax: 1 2039745790
Email: krysten.bold@yale.edu

Abstract

Background: Identifying novel ways to recruit smokers for treatment studies is important. In particular, certain subgroups of adult smokers, such as heavy-drinking smokers, are at increased risk for serious medical problems and are less likely to try quitting smoking, so drawing this hard-to-reach population into treatment is important for improving health outcomes.

Objective: This study examined the utility of Facebook advertisements to recruit smokers and heavy-drinking smokers for treatment research and evaluated smoking and alcohol use and current treatment goals among those who responded to the Web-based survey.

Methods: Using Facebook's advertising program, 3 separate advertisements ran for 2 months targeting smokers who were thinking about quitting. Advertisements were shown to adult (at least 18 years of age), English-speaking Facebook users in the greater New Haven, Connecticut, area. Participants were invited to complete a Web-based survey to determine initial eligibility for a smoking cessation research study.

Results: Advertisements generated 1781 clicks and 272 valid, completed surveys in 2 months, with one advertisement generating the most interest. Facebook advertising was highly cost-effective, averaging $0.27 per click, $1.76 per completed survey, and $4.37 per participant meeting initial screening eligibility. On average, those who completed the Web-based survey were 36.8 (SD 10.4) years old, and 65.8% (179/272) were female. Advertisements were successful in reaching smokers; all respondents reported daily smoking (mean 16.2 [SD 7.0] cigarettes per day). The majority of smokers (254/272, 93.4%) were interested in changing their smoking behavior immediately. Many smokers (161/272, 59.2%) also reported heavy alcohol consumption at least once a month. Among smokers interested in reducing their alcohol use, more were heavy drinkers (45/56, 80.4%) compared to non-heavy drinkers (11/56, 19.6%; $\chi^2_{[1,N=272]}$=13.0, P<.001). Of those who met initial screening eligibility from the Web-based survey, 12.7% (14/110) attended an in-person follow-up appointment.

Conclusions: Social media advertisements designed to target smokers were cost-effective and successful for reaching adult smokers interested in treatment. Additionally, recruiting for smokers reached those who also drink alcohol heavily, many of whom were interested in changing this behavior as well. However, additional social media strategies may be needed to engage individuals into treatment after completion of Web-based screening surveys.

(J Med Internet Res 2016;18(5):e107) doi:10.2196/jmir.5360

KEYWORDS

smoking; alcohol drinking; social media; research subject recruitment

Introduction

Identifying novel ways to recruit smokers for treatment is important. Tobacco use remains a large public health problem [1,2], and most smokers who try to quit each year do so without any form of treatment (68.3%) [3] and rarely achieve lasting abstinence [3]. Certain subgroups of adult smokers, such as heavy-drinking smokers, may be especially at risk given evidence that they are less likely to try quitting smoking [4-7] and have greater odds of serious medical problems, including cancers, liver cirrhosis, and pancreatitis [8-11]. Thus, drawing smokers and heavy-drinking smokers into treatment is important for improving health outcomes.

Research suggests there is a large subgroup of smokers who also drink heavily. In national samples, approximately 6.2 million US adults have co-occurring nicotine dependence and an alcohol use disorder [12]. Additionally, rates of binge (ie, ≥5 drinks on the same occasion) and heavy alcohol use (ie, binge use ≥5 times in the past month) are more than twice as high in smokers compared with non-smokers (binge: 42.9% smokers, 17.5% nonsmokers; heavy drinking: 15.7% smokers, 3.8% nonsmokers) [13].

Although most heavy drinkers do not seek treatment [14], smoking status may be a way of identifying heavy drinkers [15] and engaging them in treatment. Recently, Internet-based recruitment strategies have been evaluated as alternative methods for reaching and engaging individuals into treatment and clinical research. Social media advertising, such as through Facebook, was more cost-effective for reaching young adult smokers in a national Web-based survey study, totaling $4.28 per survey, and produced more valid results than other Web-based advertisements (eg, Craigslist) or recruitment methods through survey companies [16]. Another study indicated that Web-based recruitment captured a greater proportion of people with substance use and mental health issues compared with alternative recruitment methods (ie, flyers, university courses) [17]. Others have successfully used social media advertising to reach targeted audiences for substance use research, including difficult-to-recruit populations such as young adults [18], veterans [19], immigrants [20], and Latino smokers [21]. These studies suggest Web-based advertising may be a useful tool to consider for recruiting heavy-drinking smokers as well.

Given the widespread use of social media by the general adult population [22], and research highlighting the usefulness of Web-based recruitment strategies, we sought to examine the utility of Facebook advertising to recruit adult smokers from the community for treatment. The primary aims of the study were to (1) evaluate the effectiveness of Facebook advertising to reach adult smokers and a subpopulation of heavy-drinking smokers and (2) characterize smoking and alcohol use and current treatment goals among responders. Identifying who responds to smoking advertisements on social media may suggest new ideas for recruitment or innovative Web-based interventions to address substance use.

Methods

Participants

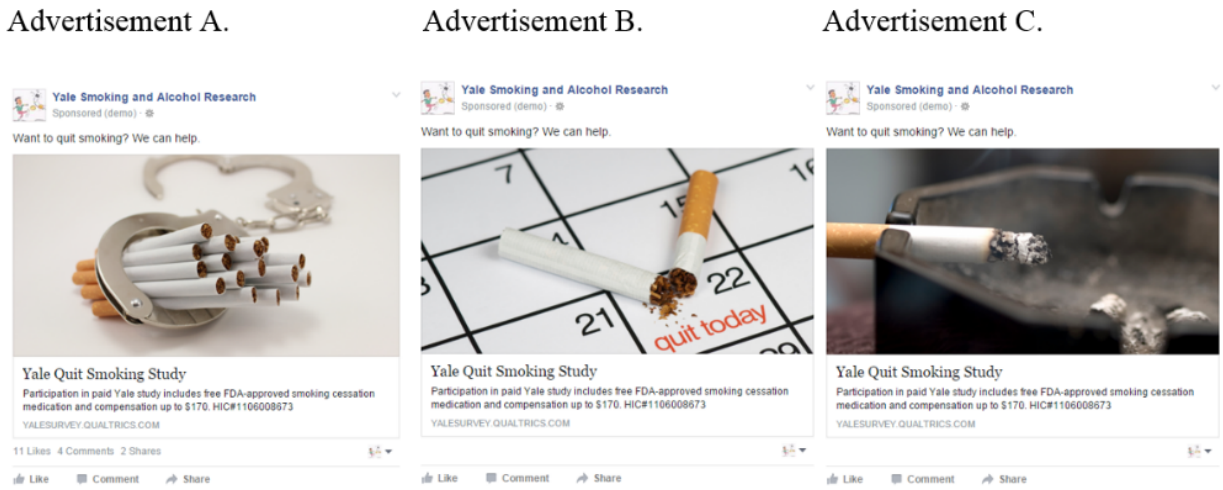
The target population was adult smokers living in the greater New Haven, Connecticut, area who were interested in participating in smoking cessation treatment research. Individuals had to be literate in English to complete the Web-based survey.

Facebook Recruitment

We utilized Facebook's advertising program for 2 months from April 13, 2015 through June 16, 2015. We created 3 advertisements that appeared on the Facebook pages of our target audience (at least 18 years old) within a geographic radius of 30-40 miles from New Haven, Connecticut, given our interest in recruiting individuals able to participate in an in-person treatment research study. All 3 advertisements targeted smokers interested in treatment by presenting images of cigarettes with the phrase "Want to quit smoking? We can help" (Figure 1). Facebook approved all advertisements. The presentation of advertisements to users was not randomized; all 3 advertisements were uploaded simultaneously to the Facebook platform with identical settings. The Facebook advertisement delivery system was set to optimize advertisement presentation over time. This system rotates through the advertisement presentation to display all 3 advertisements in the set, and when it is determined that an advertisement is performing better than others (ie, greater number of clicks), this advertisement is selected by the program to be displayed more often.

We selected days when we wanted the advertisements to run; the advertisements ran for 14 days in total over 2 months. We specified a spending limit for the entire advertising campaign (3 advertisements) of $25 per day. On days when advertisements were running, Facebook set pricing for the campaign (ie, a "bid") that was used to calculate a competitive value (ie, sum of the bid plus the advertisement's intrinsic quality, measured by past performance). Bids are entered into an auction where all advertisers compete for their advertisement to be shown to their target audience. Facebook monitors competing bids from other advertisers and optimized our advertisement pricing based on our spending limit to enhance the likelihood that the advertisement would win an auction and be visible to Facebook users [23]. Facebook monitored the number of impressions (ie, each time an advertisement is shown to a user), total reach (ie, number of people seeing the advertisements), clicks on the advertisements, and total cost for all advertisements. We used this data to evaluate the success of each advertisement and cost-effectiveness in obtaining survey responses. In accordance with Facebook's data use agreement, we present aggregated anonymous results.

Figure 1. Facebook advertisements designed to target smokers interested in quitting.



Study Procedures

The Yale Institutional Review Board approved all study procedures. After clicking on the advertisement, participants were brought to a website that described the purpose of the survey for determining preliminary eligibility for a treatment research study. Participants were informed that their responses were confidential and were told they could skip questions they did not wish to answer although this may affect their eligibility. Participants were given contact information for research staff if they had questions or preferred to complete the screening by phone. Participants provided informed consent by selecting "yes," indicating they understood this information, were at least 18 years old, and wished to complete the Web-based survey. Participants were not offered any incentives for completing the survey, but the advertisement indicated paid participation in the research study (up to $170) and free FDA-approved smoking cessation treatment if eligible. Survey participants entered contact information (email addresses and/or phone number) to facilitate scheduling an intake for the treatment research study. Responses were screened for duplicate entries and were excluded from the current analysis if the same email address or phone number was listed across multiple surveys.

Assessments

The Web-based survey contained 18 questions and took 5-10 minutes to complete. Multiple-choice responses provided an option for participants to select "No" or "I choose not to answer."

Demographics

To ensure that Web-based surveys were brief, data were only collected about participant age and sex, which were used to assess initial eligibility for the main treatment research study (eg, females who were currently pregnant or nursing were excluded).

Smoking History

Participants reported the number of cigarettes they currently smoke and their interest in reducing their cigarette use rated "immediately," "at a later date," "I am not interested in reducing my cigarette use," or "I choose not to answer."

Alcohol Use History

Frequency of heavy alcohol use was assessed by asking participants to select how frequently they consumed 4 or more (if female) or 5 or more (if male) standard drinks using National Institute on Alcohol Abuse and Alcoholism (NIAAA) guidelines for heavy drinking [24]. Categorical response options included "every day," "5 to 6 times a week," "3 to 4 times a week," "1 to 2 times a week," "2 to 3 times a month," "once a month," and "never." Participants were provided information about standard drink equivalents (ie, 12-ounce bottle of beer, 5-ounce glass of wine, one shot of hard liquor by itself or in a mixed drink). Desire to reduce alcohol use was rated "yes," "maybe at a later date," or "no." Alcohol withdrawal symptoms were assessed by asking participants if they experienced 10 specific symptoms when they cut down or stopped drinking (eg, sweating, vomiting, seizure, hallucinations) rated "yes," "no," "I choose not to answer," or "N/A" if participants had never cut down or stopped drinking.

Other Eligibility Screening Questions

Screening questions regarding whether participants would be available and eligible for the in-person treatment program included "are you currently taking part in any other research study," "do you have a permanent address and phone number (yes or no)," and "are you planning on moving out of the greater New Haven area in the next 6 months?" Additionally, participants were asked if they were currently pregnant or nursing, being treated for any medical problems, currently taking any prescription medications, and had a history of treatment for psychiatric problems. They were prompted to provide free-response text to specify the condition or medication.

Initial Screening Eligibility

Completed surveys were evaluated by research staff to determine initial eligibility for participation in a treatment research study. Eligibility criteria for being invited for an in-person intake evaluation included being at least 18 years old, currently smoking cigarettes, reporting at least 1 heavy drinking day in the past month, no history of severe alcohol withdrawal (eg, hallucinations), not currently participating in another research study, reporting having a permanent address and phone number, not planning to move in the next 6 months, not currently

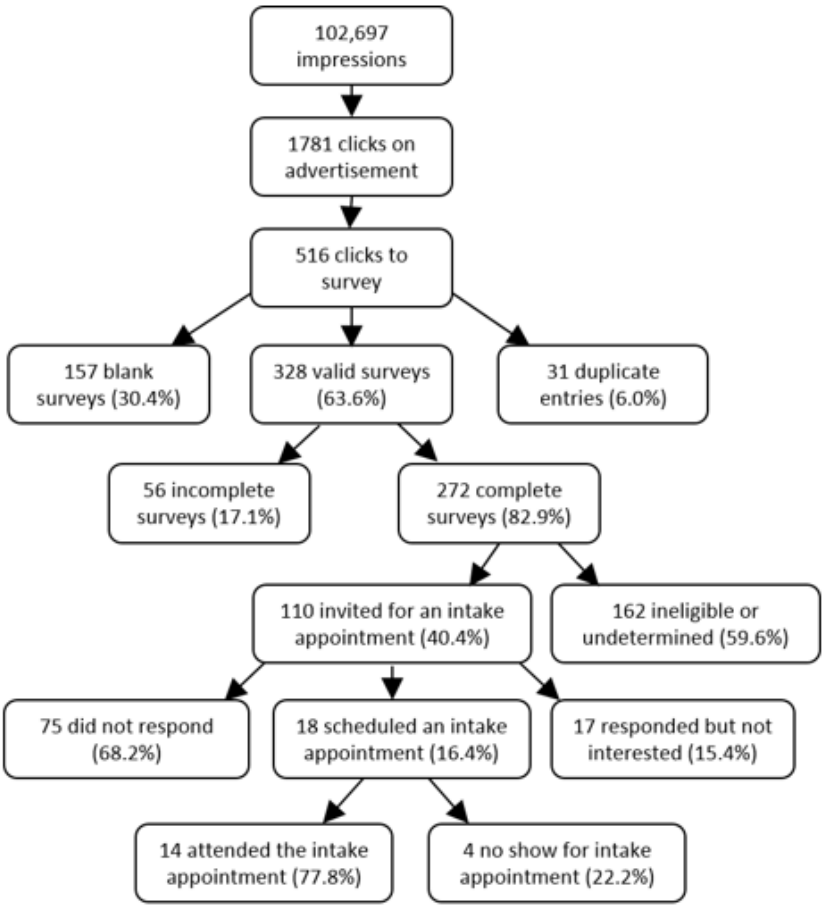
pregnant or nursing, not currently using medication to treat alcohol or tobacco use (eg, naltrexone, varenicline, bupropion, nicotine replacement), and no history of serious psychiatric illness (eg, schizophrenia, bipolar disorder).

Results

Recruitment Results

Recruitment results are depicted in Figure 2. In total, 3 Facebook advertisements generated 102,697 impressions (ie, number of times the advertisement was displayed), 1781 clicks on the advertisement, and 516 surveys in 2 months, at an overall cost of $480.89. Facebook estimated we had a daily reach (ie, number of people who saw the advertisement) of 3400-9100 out of 890,000 (the total number of people in our selected audience who were active on Facebook each day). The majority of people viewed and clicked on our advertisements from mobile devices (1662/1781, 93%, vs 113/1781, 6% on tablets and 6/1781, <1% on desktops). Of the 3 advertisements, advertisement A (Figure 1) generated the most interest and clicks to the survey page (1393/1781, 78.2% of clicks). A total of 328 valid Web-based surveys were analyzed after removing blank (n=157) and duplicate (n=31) entries. Of the 328 surveys, 55 (16.8%) answered only initial questions related to demographics and interest in changing smoking or drinking behavior. Completed surveys were reviewed, resulting in 110 participants who met preliminary screening criteria. Research staff contacted all 110 participants to let them know more about the study and invite them to complete an intake and further evaluation for a treatment research study. Of those, 18 were interested in the treatment study and scheduled an intake appointment. Most participants who scheduled an intake (14/18) attended this appointment. Overall, Facebook advertising was highly cost-effective, averaging $0.27 per click, $1.76 per completed survey, $4.37 per participant eligible for an intake, and $34.35 per participant who completed the in-person intake.

Figure 2. Flow diagram showing response rates to Facebook advertisements.



Participant Characteristics

Descriptive statistics were used to characterize participant demographics and responses to survey items. On average, survey responders were 37.4 (SD 10.8) years old, and 63.4% (208/328) were female. Independent sample *t* tests and chi-square analyses were used to assess differences in demographic characteristics by survey completer status (Table 1). Age (t_{325}=−1.86, *P*=.06), desire to reduce cigarette use ($\chi^2[_{2,N=328}]$=0.6, *P*=.73), and desire to reduce alcohol use ($\chi^2[_{2,N=328}]$=2.3, *P*=.32) did not differ significantly by completion status. Significantly more females completed the survey (179/272, 65.8%) than males (93/272, 34.2%; $\chi^2[_{1,N=328}]$=3.9, *P*=.04).

Table 1. Demographic characteristics by survey completion status.

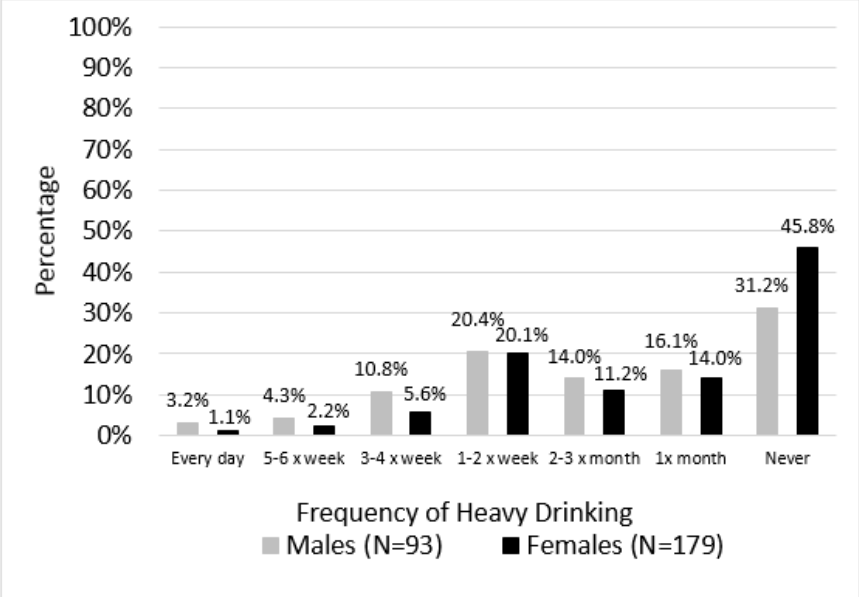
Demographic characteristics		Incomplete survey (N=56)	Complete survey (N=272)
Age in years, mean (SD)		39.8 (12.2)	36.8 (10.4)
Sex, female, n (%)		29 (51.8)	179 (65.8)
Desire to reduce cigarette use, n (%)			
	Immediately	53 (94.6)	254 (93.4)
	At a later date	3 (5.4)	15 (5.5)
	No response	0 (0.0)	3 (1.1)
Desire to reduce alcohol use, n (%)			
	Yes	15 (26.8)	56 (20.6)
	Maybe at a later date	9 (16.1)	32 (11.8)
	No	32 (57.1)	184 (67.6)

Smoking and Alcohol Use

The remaining analyses characterized responses from the 272 completed surveys. The advertisement was successful in reaching smokers interested in treatment. All survey respondents (n=272) reported daily smoking and the majority stated an interest in immediately reducing their cigarette use. Subjects reported smoking 16.2 cigarettes (SD 7.0) per day on average. Additionally, many of the smokers who responded also regularly used alcohol. More than 59% (161/272) reported heavy alcohol consumption at least once a month, measured according to NIAAA guidelines as ≥ 5 standard drinks/men, ≥ 4 standard drinks/women [24]. Figure 3 shows the frequency of heavy drinking reported by men and women in the sample. More than 19% (54/272) of respondents indicated they that experienced at least one alcohol withdrawal symptom in the past.

Although most respondents did not wish to reduce their alcohol use at the time of completing the survey, almost one-third were interested in making changes now or in the future. Chi-square analyses were used to examine whether treatment interest differed between those who reported heavy drinking at least once a month and those who reported never experiencing heavy drinking. The majority of smokers also interested in reducing their alcohol use now were heavy drinkers (45/56, 80.4%) compared to non-heavy drinkers (11/56, 19.6%; $\chi^2_{[1,N=272]}$=13.0, P<.001).

Figure 3. Frequency of heavy alcohol use by sex.



Medical and Psychiatric History

Participants responded to questions about current medical and psychiatric problems to assess initial eligibility. Overall, 1.5% (4/272) were currently pregnant or nursing, 22.1% (60/272) reported currently being treated for a medical problem, 17.3% (47/272) reported a history of treatment for psychiatric problems, 31.6% (86/272) reported currently taking medications or prescription drugs, and less than 5% (13/272) selected that they preferred not to answer these questions.

Discussion

This study evaluated the utility of Facebook advertising to recruit smokers and heavy-drinking smokers for treatment. Using 3 advertisements over 2 months, we were able to target adults interested in smoking cessation treatment, many of whom also drank heavily and were interested in changing their alcohol use as well. The majority of clicks on the survey from our advertisements resulted in valid, completed survey responses. Furthermore, the obtained information allowed for a cost-effective way of conducting initial eligibility screening, although additional efforts may be needed to engage individuals into treatment after completion of Web-based screening surveys.

Overall, Facebook was effective for reaching adult smokers and heavy-drinking smokers. However, only 12.7% of those who were initially eligible based on survey responses attended an in-person intake meeting for treatment research. The average cost per valid completed survey was comparable to costs from social media advertising reported by Ramo and colleagues [16]. Although the cost per participant completing the intake appointment was higher ($34), this was comparable or cheaper than other investigations using Web-based social media to recruit for smoking studies ($30-$170) [25-28], and was less expensive than traditional recruitment costs (eg, direct mailing or print journalism) reported in other studies ($50-$600) [26,29,30]. Only a small proportion started but did not complete the survey. We successfully obtained responses about tobacco and alcohol use for the majority of respondents who initiated the survey, suggesting Web-based surveys may be especially useful for quickly obtaining a range of information from target audiences, including information on potentially stigmatized behavior such as substance use. Additionally, elements of online communication such as convenience, familiarity, or anonymity may contribute to individuals' greater willingness to complete Web-based surveys compared with face-to-face meetings. Although this may reduce the rate of in-person follow-up from Web-based advertising and potentially limit the utility for recruiting smokers into traditional treatment in practice, it is also possible that Web-based advertising reaches individuals who would not otherwise present for treatment. Identifying more effective ways to encourage in-person follow-up or capitalize on online convenience will be important for treatment dissemination.

In particular, reaching heavy-drinking smokers through smoking cessation treatment may provide unique opportunities for motivational enhancement interventions for those who are not currently interested in treatment for their alcohol use. Although most survey respondents were interested in changing their smoking behavior, fewer were interested in changing their alcohol use. However, our results suggest there is a subset of smokers who are interested in changing both tobacco and alcohol use behaviors at the same time. These findings are consistent with prior work [31,32] and support the idea of integrating smoking cessation services into treatment for co-occurring substance use [33-36].

This study has several limitations that should be considered. First, we are unable to determine the representativeness of the sample recruited through Facebook. Data generated by Facebook do not allow us to assess how our sample of respondents compares with the general population of Facebook users meeting our target criteria. Additionally, our sample is from a limited geographic region given our interest in using Web-based recruitment to generate participants for an in-person treatment study. We only collected demographic information related to sex and age, thus we are limited in our ability to determine how representative this sample was of our geographic region or how these results may compare with other regions. However, the average age and cigarettes per day of our sample were similar to other studies using Internet-based recruitment for adult smokers [25,37], supporting the generalizability of these findings. Furthermore, our 3 advertisements were viewed an unequal number of times as designed by the Facebook algorithm (ie, more popular advertisements were shown more often), and without systematically varying the advertisement content and presentation of images, it is difficult to determine which elements of advertisement A were most appealing. It is possible that something about the advertisement itself resonated more with the audience, although additional work is needed to determine the most effective social media advertising image. Lastly, we are limited by the self-report nature of the assessments, and survey results may be influenced by reporting bias. We are unable to determine the reason for incomplete survey responses. Researchers should carefully consider participant factors or other technological factors that may affect Web-based survey completion.

Despite these limitations, our results suggest Facebook advertisements were successful in reaching adult smokers interested in treatment research. Additionally, general recruitment for smokers reached those who also drink alcohol heavily, which may provide an opportunity to engage heavy drinkers into treatment. Given the widespread use of social media and the ease of connecting to users on the go through mobile devices, Web-based advertising for survey or intervention research may be an important strategy to engage at-risk and hard-to-reach populations.

Acknowledgments

We would like to thank Kassondra Bertulis for her contribution to this project.

This research was supported by grants from the National Institutes of Health K23AA020000 (LMF), KO5014715 (SOM), T32DA019426 (KWB), and by the State of Connecticut, Department of Mental Health and Addiction Services. These sponsors had no role in the study design; collection, analysis, or interpretation of the data; writing the manuscript; or the decision to submit the paper for publication.

Conflicts of Interest

Stephanie S O'Malley is a member of the American Society of Clinical Psychopharmacology's Alcohol Clinical Trials Initiative, supported with funding from Abbott Laboratories, Alkermes, Eli Lilly, Lundbeck, Pfizer, and Ethypharma; consultant/advisory board member, Alkermes; contract as a site for a multisite study, Eli Lilly; medication supplies, Pfizer; Scientific Panel Member, Hazelden Betty Ford Foundation. All other authors report no conflicts of interest.

References

1. Centers for Disease Control and Prevention (CDC). Smoking-attributable mortality, years of potential life lost, and productivity losses--United States, 2000-2004. MMWR Morb Mortal Wkly Rep 2008 Nov 14;57(45):1226-1228 [FREE Full text] [Medline: 19008791]
2. U.S. Department of Health and Human Services. The health consequences of smoking: 50 years of progress. A report of the surgeon general. Atlanta: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health; 2014.
3. Centers for Disease Control and Prevention (CDC). Quitting smoking among adults--United States, 2001-2010. MMWR Morb Mortal Wkly Rep 2011 Nov 11;60(44):1513-1519 [FREE Full text] [Medline: 22071589]
4. Cook JW, Fucito LM, Piasecki TM, Piper ME, Schlam TR, Berg KM, et al. Relations of alcohol consumption with smoking cessation milestones and tobacco dependence. J Consult Clin Psychol 2012 Dec;80(6):1075-1085 [FREE Full text] [doi: 10.1037/a0029931] [Medline: 22963593]
5. Dawson DA. Drinking as a risk factor for sustained smoking. Drug Alcohol Depend 2000 Jun 1;59(3):235-249. [Medline: 10812284]
6. Osler M, Prescott E, Godtfredsen N, Hein HO, Schnohr P. Gender and determinants of smoking cessation: a longitudinal study. Prev Med 1999 Jul;29(1):57-62. [doi: 10.1006/pmed.1999.0510] [Medline: 10419801]
7. Zimmerman RS, Warheit GJ, Ulbrich PM, Auth JB. The relationship between alcohol use and attempts and success at smoking cessation. Addict Behav 1990;15(3):197-207. [Medline: 2378280]
8. Kuper H, Tzonou A, Kaklamani E, Hsieh CC, Lagiou P, Adami HO, et al. Tobacco smoking, alcohol consumption and their interaction in the causation of hepatocellular carcinoma. Int J Cancer 2000 Feb 15;85(4):498-502 [FREE Full text] [Medline: 10699921]
9. Lowenfels AB, Maisonneuve P, Cavallini G, Ammann RW, Lankisch PG, Andersen JR, et al. Prognosis of chronic pancreatitis: an international multicenter study. International Pancreatitis Study Group. Am J Gastroenterol 1994 Sep;89(9):1467-1471. [Medline: 8079921]
10. Vaillant GE, Schnurr PP, Baron JA, Gerber PD. A prospective study of the effects of cigarette smoking and alcohol abuse on mortality. J Gen Intern Med 1991;6(4):299-304. [Medline: 1890499]
11. Znaor A, Brennan P, Gajalakshmi V, Mathew A, Shanta V, Varghese C, et al. Independent and combined effects of tobacco smoking, chewing and alcohol drinking on the risk of oral, pharyngeal and esophageal cancers in Indian men. Int J Cancer 2003 Jul 10;105(5):681-686 [FREE Full text] [doi: 10.1002/ijc.11114] [Medline: 12740918]
12. Falk DE, Yi HY, Hiller-Sturmhöfel S. An epidemiologic analysis of co-occurring alcohol and tobacco use and disorders: findings from the National Epidemiologic Survey on Alcohol and Related Conditions. Alcohol Res Health 2006;29(3):162-171. [Medline: 17373404]
13. Substance Abuse and Mental Health Services Administration. Results from the 2013 National Survey on Drug Use and Health: Summary of National Findings. Rockville, MD: NSDUH Series H-48, HHS Publication No. (SMA). Substance Abuse and Mental Health Services Administration; 2014:14-4863.
14. Witkiewitz K, Dearing RL, Maisto SA. Alcohol use trajectories among non-treatment-seeking heavy drinkers. J Stud Alcohol Drugs 2014 May;75(3):415-422 [FREE Full text] [Medline: 24766753]
15. McKee SA, Falba T, O'Malley SS, Sindelar J, O'Connor PG. Smoking status as a clinical indicator for alcohol misuse in US adults. Arch Intern Med 2007 Apr 9;167(7):716-721 [FREE Full text] [doi: 10.1001/archinte.167.7.716] [Medline: 17420431]
16. Ramo DE, Hall SM, Prochaska JJ. Reaching young adult smokers through the internet: comparison of three recruitment mechanisms. Nicotine Tob Res 2010 Jul;12(7):768-775 [FREE Full text] [doi: 10.1093/ntr/ntq086] [Medline: 20530194]
17. Thornton LK, Harris K, Baker A, Johnson M, Kay-Lambkin FJ. Recruiting for addiction research via Facebook. Drug Alcohol Rev 2015 Jul 14;Epub ahead of print:1-9. [doi: 10.1111/dar.12305] [Medline: 26179224]
18. Ramo DE, Prochaska JJ. Broad reach and targeted recruitment using Facebook for an online survey of young adult substance use. J Med Internet Res 2012;14(1):e28 [FREE Full text] [doi: 10.2196/jmir.1878] [Medline: 22360969]
19. Pedersen ER, Helmuth ED, Marshall GN, Schell TL, PunKay M, Kurz J. Using facebook to recruit young adult veterans: online mental health research. JMIR Res Protoc 2015;4(2):e63 [FREE Full text] [doi: 10.2196/resprot.3996] [Medline: 26033209]
20. Carlini BH, Safioti L, Rue TC, Miles L. Using Internet to recruit immigrants with language and culture barriers for tobacco and alcohol use screening: a study among Brazilians. J Immigr Minor Health 2015 Apr;17(2):553-560. [doi: 10.1007/s10903-013-9934-1] [Medline: 24563138]

21. Graham AL, Fang Y, Moreno JL, Streiff SL, Villegas J, Muñoz RF, et al. Online advertising to reach and recruit Latino smokers to an internet cessation program: impact and costs. J Med Internet Res 2012;14(4):e116 [FREE Full text] [doi: 10.2196/jmir.2162] [Medline: 22954502]
22. Chou WS, Hunt YM, Beckjord EB, Moser RP, Hesse BW. Social media use in the United States: implications for health communication. J Med Internet Res 2009;11(4):e48 [FREE Full text] [doi: 10.2196/jmir.1249] [Medline: 19945947]
23. Facebook. Advertiser help center: scheduling and delivery. 2016. URL: https://www.facebook.com/business/help/430291176997542 [accessed 2016-01-22] [WebCite Cache ID 6h2FkW82n]
24. NIAAA. NIAAA Council Task Force. 2003. Recommended alcohol questions URL: http://www.niaaa.nih.gov/research/guidelines-and-resources/recommended-alcohol-questions [accessed 2015-11-16] [WebCite Cache ID 6d5XHutJJ]
25. Frandsen M, Walters J, Ferguson SG. Exploring the viability of using online social media advertising as a recruitment method for smoking cessation clinical trials. Nicotine Tob Res 2014 Feb;16(2):247-251. [doi: 10.1093/ntr/ntt157] [Medline: 24127266]
26. Rait MA, Prochaska JJ, Rubinstein ML. Recruitment of adolescents for a smoking study: use of traditional strategies and social media. Transl Behav Med 2015 Sep;5(3):254-259 [FREE Full text] [doi: 10.1007/s13142-015-0312-5] [Medline: 26327930]
27. Ramo D, Rodriguez T, Chavez K, Sommer M, Prochaska J. Facebook recruitment of young adult smokers for a cessation trial: Methods, metrics, and lessons learned. Internet Interventions 2014 Apr;1(2):58-64.
28. Heffner JL, Wyszynski CM, Comstock B, Mercer LD, Bricker J. Overcoming recruitment challenges of web-based interventions for tobacco use: the case of web-based acceptance and commitment therapy for smoking cessation. Addict Behav 2013 Oct;38(10):2473-2476 [FREE Full text] [doi: 10.1016/j.addbeh.2013.05.004] [Medline: 23770645]
29. Gordon J, Akers L, Severson H, Danaher B, Boles S. Successful participant recruitment strategies for an online smokeless tobacco cessation program. Nicotine Tob Res 2006 Dec;8(Suppl 1):S35-S41.
30. Boyle RG, Enstad C, Asche SE, Thoele MJ, Sherwood NE. Evaluating strategies and costs to recruit smokeless tobacco users. Addict Behav 2007 Dec;32(12):3088-3092. [doi: 10.1016/j.addbeh.2007.06.005] [Medline: 17602843]
31. Fucito LM, Hanrahan TH. Heavy-drinking smokers' treatment needs and preferences: A mixed-methods study. J Subst Abuse Treat 2015 Dec;59:38-44. [doi: 10.1016/j.jsat.2015.07.001] [Medline: 26297324]
32. Gulliver SB, Kamholz BW, Helstrom AW. Smoking cessation and alcohol abstinence: what do the data tell us? Alcohol Res Health 2006;29(3):208-212. [Medline: 17373411]
33. Cooney NL, Litt MD, Sevarino KA, Levy L, Kranitz LS, Sackler H, et al. Concurrent alcohol and tobacco treatment: Effect on daily process measures of alcohol relapse risk. J Consult Clin Psychol 2015 Apr;83(2):346-358. [doi: 10.1037/a0038633] [Medline: 25622198]
34. Fiore M, Jaen C, Baker T, Bailey W, Benowitz N, Curry S. Treating tobacco use and dependence: 2008 update. Clinical Practice Guideline. Rockville, MD: U.S. Department of Health and Human Services, Public Health Service; 2008.
35. Hall SM, Prochaska JJ. Treatment of smokers with co-occurring disorders: emphasis on integration in mental health and addiction treatment settings. Annu Rev Clin Psychol 2009;5:409-431 [FREE Full text] [doi: 10.1146/annurev.clinpsy.032408.153614] [Medline: 19327035]
36. Prochaska JJ, Delucchi K, Hall SM. A meta-analysis of smoking cessation interventions with individuals in substance abuse treatment or recovery. J Consult Clin Psychol 2004 Dec;72(6):1144-1156. [doi: 10.1037/0022-006X.72.6.1144] [Medline: 15612860]
37. Puckett M, Neri A, Thompson T, Underwood JM, Momin B, Kahende J, Centers for Disease Control and Prevention (CDC). Tobacco cessation among users of telephone and web-based interventions--four states, 2011-2012. MMWR Morb Mortal Wkly Rep 2015 Jan 2;63(51):1217-1221 [FREE Full text] [Medline: 25551593]

Abbreviations

NIAAA: National Institute on Alcohol Abuse and Alcoholism
FDA: Food and Drug Administration

Edited by G Eysenbach; submitted 19.11.15; peer-reviewed by E Pedersen, S Hall; comments to author 10.12.15; revised version received 02.02.16; accepted 20.02.16; published 18.05.16.

Please cite as:
Bold KW, Hanrahan TH, O'Malley SS, Fucito LM
Exploring the Utility of Web-Based Social Media Advertising to Recruit Adult Heavy-Drinking Smokers for Treatment
J Med Internet Res 2016;18(5):e107
URL: http://www.jmir.org/2016/5/e107/
doi:10.2196/jmir.5360
PMID:27194456

Original Paper

Acceptability of Interventions Delivered Online and Through Mobile Phones for People Who Experience Severe Mental Health Problems: A Systematic Review

Natalie Berry[1,2], BSc (Hons), MSc; Fiona Lobban[3], BA (Hons), PhD, DClinPsy; Richard Emsley[4], BSc (Hons), PhD; Sandra Bucci[2], BSc (Hons), ClinPsyD

[1]Health eResearch Centre (HeRC), Institute of Population Health, University of Manchester, Manchester, United Kingdom

[2]School of Psychological Sciences, Faculty of Medical and Human Sciences, University of Manchester, Manchester, United Kingdom

[3]Spectrum Centre for Mental Health, School of Health and Medicine, Lancaster University, Lancaster, United Kingdom

[4]Centre for Biostatistics, Institute of Population Health, University of Manchester, Manchester, United Kingdom

Corresponding Author:
Natalie Berry, BSc (Hons), MSc
School of Psychological Sciences
Faculty of Medical and Human Sciences
University of Manchester
Room S42, Zochonis Building
Brunswick Street
Manchester, M13 9PL
United Kingdom
Phone: 44 16130 ext 60428
Fax: 44 1613060406
Email: natalie.berry@manchester.ac.uk

Abstract

Background: Psychological interventions are recommended for people with severe mental health problems (SMI). However, barriers exist in the provision of these services and access is limited. Therefore, researchers are beginning to develop and deliver interventions online and via mobile phones. Previous research has indicated that interventions delivered in this format are acceptable for people with SMI. However, a comprehensive systematic review is needed to investigate the acceptability of online and mobile phone-delivered interventions for SMI in depth.

Objective: This systematic review aimed to 1) identify the hypothetical acceptability (acceptability prior to or without the delivery of an intervention) and actual acceptability (acceptability where an intervention was delivered) of online and mobile phone-delivered interventions for SMI, 2) investigate the impact of factors such as demographic and clinical characteristics on acceptability, and 3) identify common participant views in qualitative studies that pinpoint factors influencing acceptability.

Methods: We conducted a systematic search of the databases PubMed, Embase, PsycINFO, CINAHL, and Web of Science in April 2015, which yielded a total of 8017 search results, with 49 studies meeting the full inclusion criteria. Studies were included if they measured acceptability through participant views, module completion rates, or intervention use. Studies delivering interventions were included if the delivery method was online or via mobile phones.

Results: The hypothetical acceptability of online and mobile phone-delivered interventions for SMI was relatively low, while actual acceptability tended to be high. Hypothetical acceptability was higher for interventions delivered via text messages than by emails. The majority of studies that assessed the impact of demographic characteristics on acceptability reported no significant relationships between the two. Additionally, actual acceptability was higher when participants were provided remote online support. Common qualitative factors relating to acceptability were safety and privacy concerns, the importance of an engaging and appealing delivery format, the inclusion of peer support, computer and mobile phone literacy, technical issues, and concerns about the impact of psychological state on intervention use.

Conclusions: This systematic review provides an in-depth focus on the acceptability of online and mobile phone-delivered interventions for SMI and identified the need for further research in this area. Based on the results from this review, we recommend that researchers measure both hypothetical and actual acceptability to identify whether initial perceptions of online and mobile phone-delivered interventions change after access. In addition, more focus is needed on the potential impact of demographic and

clinical characteristics on acceptability. The review also identified issues with module completion rates and intervention use as measures of acceptability. We therefore advise researchers to obtain qualitative reports of acceptability throughout each phase of intervention development and testing. Further implications and opportunities for future research are discussed.

(J Med Internet Res 2016;18(5):e121) doi:10.2196/jmir.5250

KEYWORDS

mHealth; eHealth; severe mental health; psychosis; bipolar disorder; personality disorder; severe mental health problems (SMI); acceptability

Introduction

The exact definition of severe mental health problems (SMI) is inconsistent in the literature. Discussions have proposed that the term should be applied to describe the duration and levels of functioning in an individual, rather than focusing on specific diagnoses [1,2]. However, much of the psychological research in SMI specifically uses diagnostic criteria for participant recruitment [3]. Due to the reliance on diagnosis for recruitment in this field of research, we used a diagnostic definition of SMI for this review. Online and mobile phone-delivered interventions have the potential to improve access to evidence-based interventions for SMI; therefore, we used the Improving Access to Psychological Therapies (IAPT) initiative definition of SMI. IAPT is a United Kingdom-based National Health Service (NHS) program, which aims to increase access and availability of evidence-based psychological therapies and includes diagnoses of psychosis, bipolar disorder, and personality disorders as SMI [4]. The initiative is in line with current recommendations that people experiencing SMI be provided access to evidence-based psychological interventions, in addition to prescribed medications [5-7].

A range of barriers, including perceived stigma [8], uncertainty among practitioners about clinical effectiveness [9], cost pressures, and lack of trained facilitators [10], means that many people who could benefit from psychological interventions are often unable or unwilling to access them. Limited availability and access to psychological therapies for SMI was reflected in a survey by the charity Mind, which reported that 20% of respondents had waited for >1 year to access psychological therapies and <30% had received access within 3 months of referral [11]. In addition, a recent report by the NHS's Mental Health Taskforce highlighted that respondents' "most important" priorities for NHS mental health service improvement were early support and intervention for people experiencing SMI and increased access to psychological therapies [12]. In an attempt to reduce barriers and provide increased access to helpful interventions, researchers have investigated the role that novel technologies could play in the provision of evidence-based interventions for people with SMI.

Interventions delivered online (ie, websites) and via mobile phones (ie, smartphones, text messages, alerts, and apps) have been reported to be acceptable and show potential efficacy for the enhancement of self-care practices in individuals across a broad field, including diabetes [13,14], cancer [15], coronary heart disease [16], and psoriasis [17]. Additionally, online and mobile phone-delivered approaches have been used to implement evidence-based interventions for the promotion of health-related behaviors such as smoking cessation [18,19], physical activity [20], and weight reduction [21]. More recently, interventions have been delivered online and via mobile phones in an attempt to, among other things, reduce barriers associated with accessing mental health care [22] and to empower service users with greater choice and control over their health care needs [23]. Online and mobile phone-delivered interventions have been shown to be feasible, acceptable, and effective for depression and anxiety disorders [24,25], eating disorders [26], and substance use [27]; these interventions have since been extended to people who experience psychosis [28,29], bipolar disorder [30], and personality disorders [31], or the so-called severe mental health problems (SMI).

Levels of Internet use among people with mental health problems, including SMI, are similar to that of the general population [32]. In a survey of service users in a community psychiatric program, 59.3% of respondents reported using the Internet and 85.7% reported using mobile phones [33]. Additionally, a survey-based study for individuals with SMI reported that 72% of participants surveyed owned a mobile phone, while some expressed an interest in receiving health care services, such as appointment and medication reminders, through mobile devices [34]. A meta-analysis conducted in 2015 reported mobile phone ownership of around 81.4% in people with psychosis, reflecting similar ownership to that in the general population [35]. This meta-analysis also reported that many of the people who were surveyed expressed favorable attitudes toward mobile phone-delivered self-management strategies, for example, symptom monitoring, appointment and medication reminders, and providing an avenue for service user-provider communication. However, a more recent study of 100 individuals experiencing SMI reported high levels of traditional mobile phone ownership (85%) but lower levels of smartphone ownership (37%) [36]. The comparatively lower rates of smartphone ownership suggest that some individuals with SMI might be excluded from being able to receive interventions via smartphones; however, current levels of interest and traditional mobile phone usage in this population suggest that many would have the capabilities needed to receive online and mobile phone-delivered interventions.

Previous systematic reviews of online and mobile phone-delivered interventions for people with SMI have tended to have broad focus on acceptability, feasibility, and efficacy, rather than provide an in-depth review of one outcome. For example, a review from 2014 examined the effectiveness of online, social media, and mobile technologies for people with psychosis [28]. The authors found that interventions delivered online and via mobile devices are often feasible and acceptable,

and show potential efficacy among this group. Another review of e-mental health self-management for psychotic disorders found that individuals were willing to engage with online interventions and that such approaches can be effective for the promotion of self-management strategies [29]. More recently, a review of mobile device and eHealth interventions for people who have received a psychotic or bipolar-related diagnosis reported that interventions delivered via these modalities are both feasible and acceptable [37]. Finally, a 2015 review examined the feasibility of smartphone apps for individuals with schizophrenia [38]. The authors highlighted that the number of studies using smartphone apps for schizophrenia are limited, but they concluded that there was evidence for high feasibility due to satisfaction reports and levels of engagement. The lack of high-quality, large-scale, definitive research prevents any conclusive statements from being made. The conclusion that online and mobile phone-delivered interventions are acceptable for people with SMI has been largely based on module completion rates, intervention use, and participants' views. However, reviews have not captured the complex nature of acceptability. Specifically, reviews have not included studies investigating participant attitudes, views, and interest in interventions delivered online and via mobile phones. Additionally, potential factors that may influence acceptability, such as demographic, clinical, and intervention characteristics, have yet to be synthesized. Finally, common qualitative themes relating to the acceptability of online and mobile phone-delivered interventions for SMI have not yet been identified in systematic reviews.

In order to more closely examine acceptability, we sought to examine both the hypothetical and actual acceptability of interventions delivered online and via mobile phones for people with SMI. We define hypothetical acceptability as the acceptability of online and mobile phone-delivered interventions prior to or without an intervention being delivered, measured by participants' interest in and willingness to engage with these interventions. We define actual acceptability as the acceptability of an intervention that participants have received online or via mobile phones, which can be measured by module completion rates, intervention use, and participant views after an intervention has been delivered.

Therefore, this systematic review aimed to 1) explore whether interventions delivered online and via mobile phones are hypothetically or actually acceptable for people with SMI, 2) investigate whether participant and intervention-related factors influence acceptability, and 3) identify common participant views about acceptability from qualitative studies.

Methods

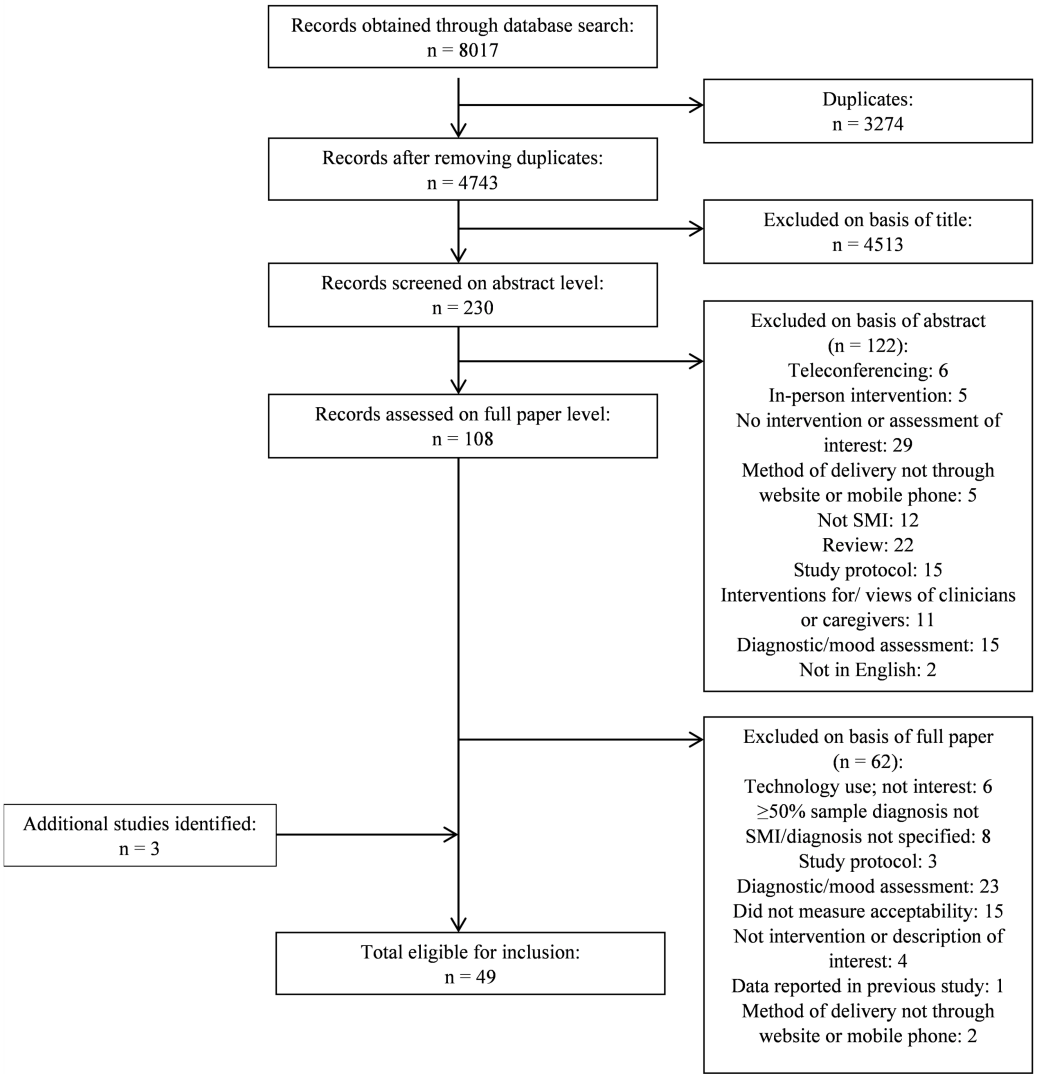
Search Strategy

This review was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) [39]. We identified studies for inclusion through searching the electronic databases PubMed, Embase, PsycINFO, CINAHL, and Web of Science. The first set of search terms were related to SMI: "psychosis" OR "psychotic" OR "psychoses" OR "schizophr*" OR "schizoaffective" OR "bipolar disorder" OR "mood disorder" OR "personality disorder" OR "severe mental illness" OR "serious mental illness" OR "severe mental health" OR "serious mental health" OR "SMI". The second set of search terms were related to online and mobile technologies: "computer" OR "technolog*", OR "digital" OR "internet" OR "online" OR "website" OR "web-based" OR "mobile" OR "phone" OR "smartphone" OR "text message" OR "SMS" OR "mHealth" OR "eHealth". These sets of search terms were linked with the Boolean operator AND.

To increase the likelihood of obtaining all of the relevant studies in the area, we took the following steps: 1) we included unpublished materials such as conference abstracts in the search; where abstracts were relevant, we contacted the lead authors for full results, 2) we used the "cited-by" function in Google Scholar to identify any eligible papers that had cited the included studies, 3) we screened references lists of included studies to gather any papers that the search terms had not identified, and 4) we contacted key authors in the area who were identified as potentially having further unpublished results.

We produced a flowchart of each stage of the database search, with the search yielding a total of 8017 papers (Figure 1).

Figure 1. Flow diagram of systematic search for interventions delivered online and through mobile phones for people with severe mental health problems (SMI).



Eligibility Criteria

We conducted full database searches in April 2015, with the inclusion and exclusion criteria identified prior to the collection period. Due to the relatively new nature of the field of the Internet and mobile phones and mental health, we considered only studies from 2005–2015. We included studies if they recruited participants with a diagnosis of psychosis, bipolar disorder, or personality disorder. Studies where these diagnoses were self-reported (not confirmed by a clinician or through an initial assessment) were still included due to Web-based recruitment strategies often used in the field. We defined interventions as a method used online or via a mobile phone with the aim of modifying a participant's behavior or psychological well-being. Therefore, we aimed to provide a broad range of interventions, from more simplistic interventions, such as medication and appointment reminders, to more complex interventions, such as cognitive behavioral therapy and interactive psychoeducation. We included studies if they involved the delivery of an intervention online or via a mobile phone or if they investigated participants' interest in and willingness to receive interventions delivered in these formats. Supported interventions (where participants were supported by a trained facilitator or where online and mobile phone-delivered interventions were implemented in conjunction with face-to-face therapy) were also included because some of the studies offered optional support and often continued face-to-face treatment as usual. Finally, we included studies if they measured acceptability in the form of module completion rates, intervention use, or participant views. Therefore, we included qualitative, quantitative, or mixed-methods study designs.

We excluded studies if the intervention was solely focused on diagnostic assessments, mood assessments, or symptom monitoring where feedback was not provided to participants. In addition, we excluded studies involving telepsychiatry, for example, video and telephone calls. We included studies that involved a combination of caregivers and service users but we report characteristics and outcomes only for the participants from the service user sample. Studies that involved mixed samples were excluded if ≤50% of the participants sampled had a diagnosis of psychosis, bipolar disorder, or personality disorder. For those where >50% of participants had 1 of these

diagnoses, we contacted the authors of these papers for separate results for participants with a diagnosis of psychosis, bipolar disorder, and personality disorders.

A key aim of this systematic review was to determine factors (eg, demographic and clinical characteristics) that could influence the acceptability of online and mobile phone-delivered interventions for SMI. Therefore, we screened included studies for information about the analysis and findings relating to potential predicators of acceptability. We considered relationships between factors and acceptability as significant if they were below the .05 level of significance. However, due to the early nature of this research, we also report associations on the 10% level.

Study Selection

We excluded studies on the title level if there was no mention of online or mobile phone-delivered interventions or mental health. The first author and an independent researcher screened the 230 paper abstracts for eligibility, with a moderate level of agreement obtained (κ=.66). This moderate level of agreement was primarily due to the first author being overly inclusive while screening abstracts due to the secondary outcome nature of acceptability. The research team resolved any disagreements until a consensus was reached about study inclusion. The first author screened full texts before the research team discussed and agreed on final papers for inclusion. We identified 3 additional studies through contacting authors and reference screening. We contacted 4 authors for clarification because their studies included participants with a diagnosis of major depressive disorder, and 3 of these authors provided additional unpublished material, which we have included in this review. The final number of included studies after screening on title, abstract, and full paper level was 49.

Results

We explored the hypothetical acceptability of online and mobile phone-delivered interventions for SMI in 7 studies [34,40-46] (1821 participants; sample size range 51–1237). The mean age of participants in studies where mean age was reported [34,40,41,44,45] (n=4) was 32.73 years (range 18.33–46). Studies were conducted in the United States [34,40,41,45] (n=3), the United Kingdom [42,46] (n=2), Canada [44] (n=1), and India [43] (n=1). Multimedia Appendix 1 details the study and participant characteristics.

The actual acceptability of online and mobile phone-delivered interventions for people with SMI was measured in 42 studies [31,40,47-86] (2226 participants; sample size range 4–311). The mean age of participants across the studies where mean age was reported [31,40,47-51,53,54,57,59,60,62-68,70-74,77-80,82-86] (n=27) was 40.61 years (range 27–48.8). The average number of intervention modules per sessions was 8.5 (range 4–20), while the average intervention duration was 17 weeks (range 2–78). The majority of the studies were conducted in the United States [31,40,60,62-64,66,68,72,73,78-81,86] (n=15) and Australia [47-49,59,70,71,74-76] (n=9), while the remainder were from Finland [50-55] (n=6), the United Kingdom [56-58,82,83] (n=5), the Netherlands [61,77,84,85] (n=4), Canada [65] (n=1), Sweden [67] (n=1), and the Czech Republic [69] (n=1). Multimedia Appendix 2 details the study and participant characteristics.

Hypothetical Acceptability

Multimedia Appendix 3 presents results for the hypothetical acceptability of online and mobile phone-delivered interventions for SMI. The results reported focused on 1) the proportion of participants who agreed that they would be interested in receiving online and mobile phone-delivered interventions, 2) the impact of demographic characteristics on hypothetical acceptability, 3) whether participant levels of interest in online and mobile phone-delivered interventions differed between email and text message delivery, and 4) geographical differences in hypothetical acceptability.

The hypothetical acceptability of obtaining support via social media websites was relatively high (75.5%), and two-thirds of participants (64%) were willing to have clinicians contact them via social media during symptom emergence [41] (A Rizvi, MA, written communication, May 2015). However, levels of interest in receiving information and support varied between social media platforms (85% YouTube; 58% Facebook; 39% Twitter) [44]. Interest in mobile phone check-ins to inform health care providers about symptoms was relatively low (14.4% to 41%) [34,40,43] (D Ben-Zeev, PhD, written communication, May 2015), while interest in receiving text messages from health care providers was moderate to high (45% to 76%) [45,46]. Additionally, the hypothetical acceptability of online and mobile phone-delivered appointment and medication reminders varied extensively between studies (26% to 92.5%) [34,40,42-45] (personal communication with D Ben-Zeev, May 2015).

Participant interest in mobile phone-delivered information about treatment, services, and psychoeducation was relatively low (31% to 48.1%) [34,40,43] (personal communication with D Ben-Zeev, May 2015). However, a 2015 study specifically recruiting young people with first episode psychosis reported that 90% of participants liked the idea of receiving information about physical and mental health online or via mobile phones [44].

Impact of Demographic Characteristics on Hypothetical Acceptability

No significant relationships were reported between interest in text message and email-delivered interventions and age, sex, race, and employment status [42,45]. However, in 1 study, there was an association at the 10% level between age and interest in receiving medication reminders via text messages (P=.06) [42]. The mean and median age of participants was <35 years in 2 studies [41,44] (personal communication with A Rizvi, May 2015) and ≥35 years in 3 studies [34,40,43,45]. Interest in online and mobile phone-delivered appointment and medication reminders was higher in studies where the average age of participants was <35 years (56.7% to 92.5%) in comparison with >35 years (26% to 58%).

Hypothetical Acceptability of Interventions Delivered via Text Messages Versus Emails

In 2 studies, participants were asked whether they would prefer to receive appointment and medication reminders, check-ins with providers, and health-related information via telephone calls, text messages, or emails. In both studies, participants stated a preference for receiving telephone calls (72.6% to 81%) in comparison with text messages (8.5% to 36%) and emails (1.5% to 21%) [34,40,43] (personal communication with D Ben-Zeev, May 2015). In 2 other studies, participants also reported a preference for medication and appointment reminders delivered via text messages (40% to 92.5%) in comparison with email (26% to 79.1%) [44,45].

Differences in Hypothetical Acceptability Geographically

A questionnaire developed for participants in the United States was adapted for use in India [34,43]. A larger proportion of participants in the United States were interested in using the Internet and mobile phones for check-ins with providers (39.5%) and appointment and medication reminders (41.5%) compared with participants in India (check-ins: 14%; medication and appointment reminders: 27%) [34,40,43] (personal communication with D Ben-Zeev, May 2015). However, interest in mobile phone-delivered health information and psychoeducation was of more interest to participants in India (44%) than to those in the United States (31%).

Actual Acceptability

Multimedia Appendix 4 presents results for the actual acceptability of online and mobile phone-delivered interventions for SMI. The results reported focused on 1) intervention satisfaction ratings reported by participants, 2) intervention use and module completion rates, 3) the impact of demographic and clinical characteristics on actual acceptability, and 4) a comparison of the actual acceptability between supported and unsupported online and mobile phone-delivered interventions.

Some studies measured the actual acceptability of online and mobile phone-delivered interventions for SMI through satisfaction ratings. The proportion of participants who were satisfied with the ease of use, perceived helpfulness, and perceived usefulness of the interventions, and were willing to recommend the intervention to others, was moderate to high (41% to 90.6%), although the majority of studies tended to report values of around 75% [40,47,52,60,66-68,77-79,85]. Other studies measured actual acceptability through satisfaction ratings on Likert scales. Ratings for overall satisfaction and perceived helpfulness and usefulness were on the upper ends of the Likert scales, while moderate to high ratings were reported for ease of use [32,50,62-65,80,84,86] (personal communication with L Warner, June 2015).

Intervention use was also included as a measure of actual acceptability. In the HORYZONS study, 60% of participants used the intervention for the full 4 weeks and 70% used it for at least 3 weeks, as measured by the number of intervention log-ins [47]. Additionally, the FOCUS system [40] was used by participants on 86.5% of the study days [60], while 71% of participants in the WEGWEIS project used every feature on the website [85]. In the SOAR study, every person who participated was reported to have engaged with the website (used the education material on at least 4 visits and contributed to the forum on at least 13 visits) [78,79]. However, rates of engagement with the MyRecoveryPlan website were relatively low after 3 weeks (program only: 9%; program plus coaching: 38%) [81]. Participant response rates to assessments and prompts ranged from 65% to 93.33% [32,62,63,72,73,86].

Actual acceptability was also measured through module completion rates and session attendance. Between 45% and 81% of participants completed at least half the modules of their assigned intervention [47,57,59,66,71,76], while 81% of participants in the Mental.Net project attended at least 3 of the 5 sessions [53]. Additionally, the average program completion rate of participants who remained in the LWB trial was 60% [83].

Impact of Demographic and Clinical Characteristics on Actual Acceptability

Only 3 studies investigated whether demographic characteristics influenced actual acceptability. Of these studies, 2 reported no significant relationships between actual acceptability and age, sex, or education level [54,63]. However, 1 study reported significantly higher workbook completion rates by participants who were female and older, although this was not found for educational attainment [75,76].

Several studies explored the relationship between participant psychological state and actual acceptability. The majority of these studies reported that the presence and severity of symptoms associated with SMI, for example, anxiety, depression, mania, and reduced cognitive functioning, did not predict participant satisfaction, module completion rates, and intervention use [54,60,63,75,76,86]. However, in 1 study, there was a positive association at the 10% level between completion rates and depressive symptoms (P=.06) and perceived helpfulness and baseline manic symptoms (P=.07) [78]. Additionally, there was a significant positive relationship between participant access and use of the SOAR website and severity of positive symptoms [79].

Actual Acceptability of Supported Versus Unsupported Internet-Delivered Interventions

Comparisons of actual acceptability between supported and unsupported Internet-delivered interventions for SMI were explored in 2 studies. In the MyRecoveryPlan and BEP studies, participants were assigned to receive access to either the basic programs without support or the programs plus remote coaching and support via email to help participants when using the intervention [75,76,81]. Participant completion rates in the BEP study and usage of the MyRecoveryPlan website were significantly higher in the remote support conditions than in the unsupported conditions (P <.05).

Common Qualitative Themes

Some projects included qualitative studies to investigate participant views about the acceptability of online and mobile phone-delivered interventions for SMI and ideas for future developments to increase acceptability [49,51,53-56,58,61,62,69,70,74,75,82,84,85] (Multimedia

Appendix 5). To aid conceptual understanding, we collected common participant quotes relating to acceptability by screening the papers and creating a set of key themes that emerged. The key common themes identified were 1) concerns about the safety, privacy, and security of online and mobile phone-delivered interventions, 2) the importance of an engaging and appealing intervention delivery format, 3) participants' desire for the inclusion of remote peer support, 4) individual differences in computer and mobile phone literacy and technical issues as potential barriers to acceptability, and 5) the potential impact of psychological state on motivation to engage.

Participant Safety

Participant perceptions about the safety, confidentiality, and privacy of online and mobile phone-delivered interventions were noted in several studies [48,49,54,56,58,61,62]. Participants in the HORYZONS project felt the social networking component was safe and confidential due to its anonymous nature, restricted access, and expert moderation [48,49]. However, concerns about confidentiality were raised in Beating Bipolar focus groups, and 2 participants in the randomized controlled trial felt their privacy had been compromised through accessing the program on public computers [56,58]. Focus groups in the PCR project also revealed privacy and security concerns [61], while some participants using Mental.Net were worried about the confidentiality of the computerized delivery method [54]. Finally, the pilot trial for PRISM revealed that some participants were concerned about what they would say if other people asked what the device was for [62].

Engaging and Appealing

Comments relating to the appearance and layout of online and mobile phone-delivered interventions for SMI were evident in the included studies [40,49,51,55,56,70,74,82,84]. Many participants were satisfied with the clarity, layout, and appearance of the Mental.Net, Beating Bipolar, and WEGWEIS websites [55,56,84]. The inclusion of interactive components was generally popular. Participants in the HORYZONS project felt the interactive modules were fun; however, some were overwhelmed by the number of components [49]. Participants valued the interactive components and video and audio features of Beating Bipolar [56] and ORBIT [74], and recommended more interactive and video features on the Mental.Net and WEGWEIS websites [51,55,84]. The visual aids and pictures in FOCUS received positive comments [40], while participants liked the interactive mood monitor and flash objects on MoodSwings [70]. Finally, focus group discussions for LWB revealed that the idea of an online intervention was popular due its potential for interactive elements [82].

Peer Support

Online and mobile phone-delivered interventions for SMI sometimes include the opportunity for remote peer-to-peer communication. Participants in the HORYZONS project expressed that they liked the inclusion of social networking [48,49]; however, participants using Beating Bipolar were disappointed with the lack of discussion forum activity [58]. In the modification stage of the Mental.Net website, participants suggested the provision of a discussion forum [55], while participants in the LWB focus groups requested the option to communicate with peers online [82].

Computer and Mobile Phone Literacy and Technical Issues

Due to the technical nature of online and mobile phone-delivered interventions, some participants with low technology literacy may not find interventions delivered in this format acceptable. A small number of participants using Beating Bipolar reported low engagement due to poor computer literacy [58], while nurses in the Mental.Net project noted that disruptions were often related to insufficient technology skills [54].

Issues relating to technical functioning were also raised as potentially affecting acceptability. Nurses in the Mental.Net project felt some sessions were disrupted by technical problems such as network access [54], while participants in the modification stage reported technical issues such as inactive links [55]. Some components in the PCR study did not function adequately [61], while LWB focus group participants revealed a reluctance to engage with a website containing technical errors [82].

Impact of Psychological State on Engagement

Concerns were also raised about the influence of psychological state on intervention engagement. Some participants involved in the development of Beating Bipolar and LWB expressed concerns that people experiencing severe symptoms would not be able to engage with it [56,82]. In the Beating Bipolar testing phase, some participants felt their engagement was reduced by low mood, while others felt that low mood increased their engagement [58]. Participants who did not complete the BEP believed this was due to experiencing acute symptoms, which left them unable to engage [75].

Discussion

The aim of this review was to explore both the hypothetical and actual acceptability of interventions delivered online and via mobile phones in SMI. The results support the assertion made in previous reviews that the actual acceptability of online and mobile phone-delivered interventions for SMI is relatively high [28,29,37,38]. However, unique to this review are the findings demonstrating the hypothetical acceptability of online and mobile phone-delivered interventions for SMI, potential predictors of acceptability, and qualitative themes relating to acceptability. This review also identified that acceptability is far more complex than just module completion rates and intervention use, suggesting the need for continued service user involvement and the inclusion of participant satisfaction ratings and qualitative interviews to measure acceptability.

The hypothetical acceptability of online and mobile phone-delivered interventions for SMI generally varied between studies. However, the relatively low levels of participant interest in online and mobile phone-delivered interventions evident in some of the studies indicate that some people with SMI may be negatively predisposed toward these delivery formats. This

contrasts greatly with the relatively high levels of actual acceptability observed in the included studies.

The results from this review also indicated that hypothetical acceptability was higher for interventions delivered via mobile phones than for online formats. It is, however, important to consider these findings in the context of intervention types proposed. Specifically, studies measuring hypothetical acceptability primarily examined participant interest in using mobile phone and online resources to communicate with health care providers and receive appointment and medication reminders. Due to the transportable and immediately accessible nature of mobile phones, it is unsurprising that participants stated a preference for mobile phone delivery over online. However, the delivery preferences of more complex, time-consuming, and interactive interventions such as cognitive behavioral therapy and psychoeducation remains unknown. Many people are now able to access online content with relative ease on their mobile phones; therefore, it is questionable whether the Internet and mobile phones can be viewed as separate methods of intervention delivery.

The findings also revealed that few studies investigated whether demographic and clinical characteristics were predictors of acceptability. The studies that did investigate the influence of demographic and clinical characteristics on acceptability reported no significant relationships [42,45,54,60,63,75,76,86]; however, a few studies did report a significant relationship [75,76,78,79], or an association at the 10% level [42,86]. In addition, hypothetical acceptability was higher in studies where the mean age of participants was <35 years. The varied findings and limited number of studies prevent us from drawing overall conclusions, and further research is warranted to investigate whether demographic and clinical characteristics predict acceptability.

The review findings also indicated that actual acceptability was higher for participants who were offered remote online support than for those who were not supported [75,76,81]. While it is still too early to know whether acceptability is higher in supported interventions than in unsupported interventions, our findings indicate that the provision of remote support is likely to predict acceptability. While we acknowledge that one of the key advantages of online and mobile phone-delivered interventions is the potential reduction in the cost of trained clinicians, these findings suggest that remote support could be offered to help increase the acceptability of these approaches.

We were able to compare hypothetical acceptability across different geographical areas between 2 studies [34,40,43] (personal communication with D Ben-Zeev, May 2015). Participants in the United States were more interested in health care provider check-ins and appointment and medication reminders than were participants in India. However, interest in receiving psychoeducation and service information was higher for participants in India than for those in the United States. The authors noted that differences between participant interest may reflect the increased availability of mental health information resources already available in the United States [43]. Due to the notable increase in Internet and mobile phone access in developing countries and papers reporting the potential benefits of online and mobile phone-delivered interventions in these nations, research in this area will likely increase at a fast pace over the coming years [86-89].

Key common themes identified in the qualitative studies revealed that some participants found the safety and privacy of online and mobile phone-delivered interventions for SMI acceptable [48,49], while others were concerned that confidentiality may be compromised [54,56,58,61]. Many participants felt that online and mobile phone-delivered interventions needed interactive components to increase acceptability and were generally positive about the provision of online peer support [48,49,51,55,82,84]. Based on participant feedback, it is advisable for researchers to incorporate interactive features and social networking components within online and mobile phone-delivered interventions for SMI. A few concerns about technology literacy and technical issues were reported; therefore, researchers should ensure that participants are comfortable with the chosen format for delivery and that the delivery method functions well [54,55,58,61]. The qualitative studies also revealed participant concerns that some people with SMI may struggle to engage with online and mobile phone-delivered interventions while experiencing acute symptoms [56,82]. Researchers should be mindful about the potential influence of psychological state on acceptability across the phases of illness.

Strengths and Limitations

This review had several notable strengths. First, the range of databases we searched and the list of search terms we created were comprehensive, which ensured that we obtained eligible studies in the field. Second, the review included studies with quantitative, qualitative, and mixed-methods designs, thus enabling a broad and in-depth analysis of the current work in the field. Third, studies were coextracted on the abstract level by a researcher independent of the research team to ensure eligibility criteria were accurate. In addition, our extraction of the data was systematic and we contacted authors if we required any further information.

Findings from the studies in this review should be considered in the context of some limitations. Many of the studies reviewed measured actual acceptability through module completion rates and intervention use. The sole use of these measures of acceptability is problematic due to the potential influence of other factors, for example, the number of modules available to complete, the intervention duration, financial incentives for high completion rates, technical issues with the interventions delivered, participants' engaging in other activities while logged in, and time pressures preventing engagement. Therefore, module completion rates and intervention use are unlikely to be robust direct measures of acceptability. Across studies, module completion rates and intervention use were also reported in different ways, for example, some reported the average number of modules completed, while others reported the average duration participants spent accessing an intervention. It is impossible to determine what value constitutes an "acceptable" intervention without a universal measurement applied. Future research is needed to develop more accurate ways to assess acceptability.

The majority of the studies we reviewed that measured hypothetical acceptability asked participants about their general interest and willingness to use interventions delivered online and via mobile phones. However, participants were not asked about their interest in receiving these interventions online or via mobile phones in comparison with, or in addition to, face-to-face delivery. Had these questions been phrased differently, overall hypothetical acceptability may have been very different. The review findings also highlight the very limited amount of relevant information regarding predictors of acceptability being reported in studies. We could not draw conclusions about the influence of demographic and clinical characteristics on both hypothetical and actual acceptability.

Issues relating to the heterogeneity of approaches for participant recruitment are an important consideration for the recruitment strategies in the field. Specifically, it could be argued that online recruitment methods may bias the sample toward favorable attitudes toward online and mobile phone-delivered interventions, thus increasing levels of acceptability. However, potential bias toward the acceptability of interventions delivered via these modalities may also be prevalent in more traditional routes of recruitment (ie, through service providers and clinicians). While different recruitment strategies may attract different samples within the population, broadening intervention choice and examining acceptability remain important considerations within the samples identified, regardless of the recruitment method that is used.

There were also some limitations to the method of analysis we used. First, the review excluded papers published in a language other than English. Second, some studies included both participants with SMI and participants with other mental health problems such as depression. To combat this issue, we excluded studies if ≤50% of the total study sample had SMI, so we may have missed relevant findings for those who did experience SMI.

Implications and Future Research

Although this review highlights the relatively high acceptability of online and mobile phone-delivered interventions for SMI, it also demonstrates the complex nature of acceptability and the need for continued focus in this area. A recent systematic review concluded that, rather than concentrating on acceptability, researchers should instead investigate whether online and mobile phone-delivered interventions are effective [37]. While efficacy is undoubtedly important, we argue that acceptability remains equally important because, ultimately, if an intervention is not acceptable to service users and relevant stakeholders, people are unlikely to engage with the approach, thereby directly affecting efficacy. Based on the findings in this review, we recommend that research groups measure hypothetical and actual acceptability of online and mobile phone-delivered interventions, with an increased focus on the factors that could influence acceptability.

There were different reporting styles evident in studies detailing intervention usage, module completion rates, and session attendance. For example, some studies reported the overall proportion of participants who used the intervention over the whole study period, while others reported proportions over specific weeks during the study period. Therefore, it was not possible to investigate how the acceptability of online and mobile phone-delivered interventions changed over time. Future research should explore and report how intervention usage, module completion rates, and session attendance change throughout intervention delivery to determine whether acceptability of interventions delivered via these modalities changes across time points.

Module completion rates and intervention use may be indirectly related to actual acceptability; however, participant satisfaction ratings and qualitative views provide rich data about acceptability. These rich data can be used to develop and refine online and mobile phone-delivered interventions in order to improve the overall acceptability for people with SMI. The need to use qualitative analysis to inform the design and development of interventions delivered online and via mobile phones has been recognized in protocols for future studies in the field [90-92]. It is recommended that, if researchers choose to measure acceptability through module completion rates and intervention use, satisfaction ratings and qualitative interviews be conducted to obtain the rich information needed to identify intervention acceptability. In addition, the use of qualitative interviews examining what participants feel they have actually gained from an intervention may help to highlight the specific areas of their lives that they feel may have been improved by participating in the intervention.

This review also found that hypothetical acceptability tended to be low or varied, while actual acceptability tended to be high, indicating that people with SMI may be initially reluctant to engage with online and mobile phone-delivered interventions. However, hypothetical acceptability results are largely limited and not necessarily directly comparable with the interventions being delivered. Issues with comparing hypothetical versus actual acceptability are primarily due to discrepancies in the types of interventions being explored. Specifically, studies investigating hypothetical acceptability investigated interest in using the Internet or mobile phones to facilitate health care provider contact and medication and appointment reminders, whereas studies investigating actual acceptability tended to implement more complex interventions such as cognitive behavioral therapy and psychoeducation. It is also likely that the sample of people who were asked about actual acceptability were already hypothetically open to the idea of receiving online and mobile phone-delivered interventions. We therefore suggest that researchers measure acceptability both before (hypothetical acceptability) and after (actual acceptability) an intervention is delivered. These measurements will aid the comparison of hypothetical versus actual acceptability without the limitations associated with separate samples.

Crucially, this review showed that the majority of the studies reviewed recruited participants who were already in contact with mental health services. One of the potential advantages of the Internet and mobile phones is that they could improve access to evidence-based interventions for people who are not receiving support but who need or want it. Therefore, it is important to investigate whether people who could potentially benefit the most from online and mobile phone-delivered interventions actually find these delivery formats acceptable.

Clearly, interventions delivered online and via mobile phones do have their place in the provision of self-care for people with SMI. However, research has yet to identify predictors of acceptability and whether people who are not engaged with services also find online and mobile phone-delivered approaches acceptable. The measurement of both hypothetical and actual acceptability in future studies would enable the investigation of the impact of prior expectations on acceptability and potential changes in acceptability after access. In order to obtain rich data about acceptability, we recommend the measurement of satisfaction ratings and participant views and the continued involvement of service users throughout all aspects of intervention development and delivery.

Acknowledgments

The authors would like to thank Rebecca Harrop for assisting with coextraction of the studies. The authors would also like to thank Dror Ben-Zeev, Cheryl Forchuk, and Michael Birnbaum for providing additional results. NB is a PhD student funded by the MRC Health eResearch Centre (HeRC), Farr Institute, UK (MR/K006665/1).

Authors' Contributions

All authors contributed to the development of the review. NB completed the first draft of the manuscript, with SB, FL, and RE providing critical feedback and revisions. All authors approved and contributed to the final manuscript.

Conflicts of Interest

SB and RE are investigators for a mobile-delivered intervention for early psychosis (Actissist). FL is an investigator for a Web-based program for bipolar disorder.

Multimedia Appendix 1

Characteristics of studies that measured the hypothetical acceptability of online and mobile phone-delivered interventions for severe mental health problems.

[PDF File (Adobe PDF File), 30KB - jmir_v18i5e121_app1.pdf]

Multimedia Appendix 2

Characteristics of studies that measured the actual acceptability of online and mobile phone-delivered interventions for SMI.

[PDF File (Adobe PDF File), 68KB - jmir_v18i5e121_app2.pdf]

Multimedia Appendix 3

Results and predictors relating to the hypothetical acceptability of online and mobile phone-delivered interventions for people with severe mental health problems.

[PDF File (Adobe PDF File), 34KB - jmir_v18i5e121_app3.pdf]

Multimedia Appendix 4

Results and predictors relating to the actual acceptability of online and mobile phone-delivered interventions for people with SMI measured by intervention use, module completion rates, and participant satisfaction.

[PDF File (Adobe PDF File), 55KB - jmir_v18i5e121_app4.pdf]

Multimedia Appendix 5

Participant views and example quotes from studies that measured the acceptability of online and mobile phone-delivered interventions for people with SMI.

[PDF File (Adobe PDF File), 44KB - jmir_v18i5e121_app5.pdf]

References

1. Parabiaghi A, Bonetto C, Ruggeri M, Lasalvia A, Leese M. Severe and persistent mental illness: a useful definition for prioritizing community-based mental health service interventions. Soc Psychiatry Psychiatr Epidemiol 2006 Jun;41(6):457-463. [doi: 10.1007/s00127-006-0048-0] [Medline: 16565917]
2. Ruggeri M, Leese M, Thornicroft G, Bisoffi G, Tansella M. Definition and prevalence of severe and persistent mental illness. Br J Psychiatry 2000 Aug;177:149-155 [FREE Full text] [Medline: 11026955]

3. Cooke A. Understanding Psychosis and Schizophrenia: Why People Sometimes Hear Voices, Believe Things That Others Find Strange, or Appear Out of Touch With Reality, and What Can Help. 2014. URL: https://www.bps.org.uk/system/files/user-files/Division%20of%20Clinical%20Psychology/public/understanding_psychosis_-_final_19th_nov_2014.pdf [accessed 2016-03-08] [WebCite Cache ID 6frLUP3yv]
4. Improving Access to Psychological Therapies. Severe Mental Illness. London, UK: NHS; 2015. URL: http://www.iapt.nhs.uk/smi-/ [accessed 2015-06-01] [WebCite Cache ID 6YyH9keQ2]
5. National Institute for Health and Care Excellence. Borderline Personality Disorder: Treatment and Management. London, UK: NICE; 2009. URL: https://www.nice.org.uk/guidance/CG78/chapter/1-Guidance [accessed 2015-06-01] [WebCite Cache ID 6YyJzmlu2]
6. National Institute for Health and Care Excellence. Psychosis and Schizophrenia in Adults: Treatment and Management. London, UK: NICE; 2014. URL: http://www.nice.org.uk/guidance/cg178/chapter/1-recommendations [accessed 2015-06-01] [WebCite Cache ID 6YyKK3TI7]
7. National Institute for Health and Care Excellence. Bipolar Disorder: The Assessment and Management of Bipolar Disorder in Adults, Children and Young People in Primary and Secondary Care. London, UK: NICE; 2014. URL: http://www.nice.org.uk/guidance/cg185/chapter/1-recommendations [accessed 2015-06-01] [WebCite Cache ID 6YyKbjJOG]
8. Clement S, Schauman O, Graham T, Maggioni F, Evans-Lacko S, Bezborodovs N, et al. What is the impact of mental health-related stigma on help-seeking? A systematic review of quantitative and qualitative studies. Psychol Med 2015 Jan;45(1):11-27. [doi: 10.1017/S0033291714000129] [Medline: 24569086]
9. Morriss R. Implementing clinical guidelines for bipolar disorder. Psychol Psychother 2008 Dec;81(Pt 4):437-458. [doi: 10.1348/147608308X278105] [Medline: 18215339]
10. Berry K, Haddock G. The implementation of the NICE guidelines for schizophrenia: barriers to the implementation of psychological interventions and recommendations for the future. Psychol Psychother 2008 Dec;81(Pt 4):419-436. [doi: 10.1348/147608308X329540] [Medline: 18652741]
11. Mind. We Still Need to Talk: A Report on Access to Talking Therapies. 2013 Nov. URL: http://www.mind.org.uk/media/494424/we-still-need-to-talk_report.pdf [accessed 2016-03-08] [WebCite Cache ID 6frMCs4Oc]
12. The Five Year Forward View Mental Health Taskforce: Public Engagement Findings. London, UK: NHS England; 2015 Sep. URL: https://www.england.nhs.uk/mentalhealth/wp-content/uploads/sites/29/2015/09/fyfv-mental-hlth-taskforce.pdf [accessed 2016-05-09] [WebCite Cache ID 6heNxG8PH]
13. Cotter AP, Durant N, Agne AA, Cherrington AL. Internet interventions to support lifestyle modification for diabetes management: a systematic review of the evidence. J Diabetes Complications 2014;28(2):243-251 [FREE Full text] [doi: 10.1016/j.jdiacomp.2013.07.003] [Medline: 24332469]
14. Holtz B, Lauckner C. Diabetes management via mobile phones: a systematic review. Telemed J E Health 2012 Apr;18(3):175-184. [doi: 10.1089/tmj.2011.0119] [Medline: 22356525]
15. 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):321-330 [FREE Full text] [doi: 10.1097/NCC.0000000000000159] [Medline: 24945270]
16. Park LG, Howie-Esquivel J, Chung ML, Dracup K. A text messaging intervention to promote medication adherence for patients with coronary heart disease: a randomized controlled trial. Patient Educ Couns 2014 Feb;94(2):261-268. [doi: 10.1016/j.pec.2013.10.027] [Medline: 24321403]
17. Bundy C, Pinder B, Bucci S, Reeves D, Griffiths CEM, Tarrier N. A novel, web-based, psychological intervention for people with psoriasis: the electronic Targeted Intervention for Psoriasis (eTIPs) study. Br J Dermatol 2013 Aug;169(2):329-336. [doi: 10.1111/bjd.12350] [Medline: 23551271]
18. Civljak M, Stead LF, Hartmann-Boyce J, Sheikh A, Car J. Internet-based interventions for smoking cessation. Cochrane Database Syst Rev 2013;7:CD007078. [doi: 10.1002/14651858.CD007078.pub4] [Medline: 23839868]
19. Whittaker R, McRobbie H, Bullen C, Borland R, Rodgers A, Gu Y. Mobile phone-based interventions for smoking cessation. Cochrane Database Syst Rev 2012;11:CD006611. [doi: 10.1002/14651858.CD006611.pub3] [Medline: 23152238]
20. Hamel LM, Robbins LB, Wilbur J. Computer- and web-based interventions to increase preadolescent and adolescent physical activity: a systematic review. J Adv Nurs 2011 Feb;67(2):251-268. [doi: 10.1111/j.1365-2648.2010.05493.x] [Medline: 21198800]
21. Stephens J, Allen J. Mobile phone interventions to increase physical activity and reduce weight: a systematic review. J Cardiovasc Nurs 2013;28(4):320-329 [FREE Full text] [doi: 10.1097/JCN.0b013e318250a3e7] [Medline: 22635061]
22. Lal S, Adair CE. E-mental health: a rapid review of the literature. Psychiatr Serv 2014 Jan 1;65(1):24-32. [doi: 10.1176/appi.ps.201300009] [Medline: 24081188]
23. Hollis C, Morriss R, Martin J, Amani S, Cotton R, Denis M, et al. Technological innovations in mental healthcare: harnessing the digital revolution. Br J Psychiatry 2015 Apr;206(4):263-265. [doi: 10.1192/bjp.bp.113.142612] [Medline: 25833865]
24. 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]

25. 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]
26. Schlegl S, Bürger C, Schmidt L, Herbst N, Voderholzer U. The potential of technology-based psychological interventions for anorexia and bulimia nervosa: a systematic review and recommendations for future research. J Med Internet Res 2015;17(3):e85 [FREE Full text] [doi: 10.2196/jmir.3554] [Medline: 25840591]
27. Marsch LA, Carroll KM, Kiluk BD. Technology-based interventions for the treatment and recovery management of substance use disorders: a JSAT special issue. J Subst Abuse Treat 2014 Jan;46(1):1-4 [FREE Full text] [doi: 10.1016/j.jsat.2013.08.010] [Medline: 24041749]
28. Alvarez-Jimenez M, Alcazar-Corcoles MA, González-Blanch C, Bendall S, McGorry PD, Gleeson JF. Online, social media and mobile technologies for psychosis treatment: a systematic review on novel user-led interventions. Schizophr Res 2014 Jun;156(1):96-106. [doi: 10.1016/j.schres.2014.03.021] [Medline: 24746468]
29. van der Krieke L, Wunderink L, Emerencia AC, de Jonge P, Sytema S. E-mental health self-management for psychotic disorders: state of the art and future perspectives. Psychiatr Serv 2014 Jan 1;65(1):33-49. [doi: 10.1176/appi.ps.201300050] [Medline: 24129842]
30. Hidalgo-Mazzei D, Mateu A, Reinares M, Matic A, Vieta E, Colom F. Internet-based psychological interventions for bipolar disorder: Review of the present and insights into the future. J Affect Disord 2015 Aug 28;188:1-13. [doi: 10.1016/j.jad.2015.08.005] [Medline: 26342885]
31. Rizvi SL, Dimeff LA, Skutch J, Carroll D, Linehan MM. A pilot study of the DBT coach: an interactive mobile phone application for individuals with borderline personality disorder and substance use disorder. Behav Ther 2011 Dec;42(4):589-600. [doi: 10.1016/j.beth.2011.01.003] [Medline: 22035988]
32. Trefflich F, Kalckreuth S, Mergl R, Rummel-Kluge C. Psychiatric patients' internet use corresponds to the internet use of the general public. Psychiatry Res 2015 Mar 30;226(1):136-141. [doi: 10.1016/j.psychres.2014.12.037] [Medline: 25623020]
33. Carras MC, Mojtabai R, Furr-Holden CD, Eaton W, Cullen BA. Use of mobile phones, computers and internet among clients of an inner-city community psychiatric clinic. J Psychiatr Pract 2014 Mar;20(2):94-103. [doi: 10.1097/01.pra.0000445244.08307.84] [Medline: 24638044]
34. Ben-Zeev D, Davis KE, Kaiser S, Krzsos I, Drake RE. Mobile technologies among people with serious mental illness: opportunities for future services. Adm Policy Ment Health 2013 Jul;40(4):340-343 [FREE Full text] [doi: 10.1007/s10488-012-0424-x] [Medline: 22648635]
35. Firth J, Cotter J, Torous J, Bucci S, Firth JA, Yung AR. Mobile Phone Ownership and Endorsement of “mHealth” Among People With Psychosis: A Meta-analysis of Cross-sectional Studies. Schizophr Bull 2016;42(2):448-455. [doi: 10.1093/schbul/sbv132] [Medline: 26400871]
36. Glick G, Druss B, Pina J, Lally C, Conde M. Use of mobile technology in a community mental health setting. J Telemed Telecare 2015 Oct 29. [doi: 10.1177/1357633X15613236] [Medline: 26519378]
37. Naslund JA, Marsch LA, McHugo GJ, Bartels SJ. Emerging mHealth and eHealth interventions for serious mental illness: a review of the literature. J Ment Health 2015 May 28;24(5):1-12. [doi: 10.3109/09638237.2015.1019054] [Medline: 26017625]
38. Firth J, Torous J. Smartphone apps for schizophrenia: a systematic review. JMIR Mhealth Uhealth 2015;3(4):e102 [FREE Full text] [doi: 10.2196/mhealth.4930] [Medline: 26546039]
39. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. BMJ 2009;339:b2535 [FREE Full text] [Medline: 19622551]
40. Ben-Zeev D, Kaiser SM, Brenner CJ, Begale M, Duffecy J, Mohr DC. Development and usability testing of FOCUS: a smartphone system for self-management of schizophrenia. Psychiatr Rehabil J 2013 Dec;36(4):289-296 [FREE Full text] [doi: 10.1037/prj0000019] [Medline: 24015913]
41. Birnbaum ML, Rizvi AF, Correll CU, Kane JM. Role of social media and the Internet in pathways to care for adolescents and young adults with psychotic disorders and non-psychotic mood disorders. Early Interv Psychiatry 2015 Mar 23. [doi: 10.1111/eip.12237] [Medline: 25808317]
42. Bogart K, Wong SK, Lewis C, Akenzua A, Hayes D, Prountzos A, et al. Mobile phone text message reminders of antipsychotic medication: is it time and who should receive them? A cross-sectional trust-wide survey of psychiatric inpatients. BMC Psychiatry 2014;14:15 [FREE Full text] [doi: 10.1186/1471-244X-14-15] [Medline: 24447428]
43. Jain N, Singh H, Koolwal GD, Kumar S, Gupta A. Opportunities and barriers in service delivery through mobile phones (mHealth) for severe mental illnesses in Rajasthan, India: a multi-site study. Asian J Psychiatr 2015 Apr;14:31-35. [doi: 10.1016/j.ajp.2015.01.008] [Medline: 25701069]
44. Lal S, Dell'Elce J, Tucci N, Fuhrer R, Tamblyn R, Malla A. Preferences of young adults with first-episode psychosis for receiving specialized mental health services using technology: a survey study. JMIR Mental Health 2015 May 20;2(2):e18. [doi: 10.2196/mental.4400]
45. Miller BJ, Stewart A, Schrimsher J, Peeples D, Buckley PF. How connected are people with schizophrenia? Cell phone, computer, email, and social media use. Psychiatry Res 2015 Feb 28;225(3):458-463. [doi: 10.1016/j.psychres.2014.11.067] [Medline: 25563669]

46. Sanghara H, Kravariti E, Jakobsen H, Okocha C. Using short message services in mental health services: assessing feasibility. Mental Health Rev J 2010 Jul 06;15(2):28-33. [doi: 10.5042/mhrj.2010.0369]
47. Alvarez-Jimenez M, Bendall S, Lederman R, Wadley G, Chinnery G, Vargas S, et al. On the HORYZON: moderated online social therapy for long-term recovery in first episode psychosis. Schizophr Res 2013 Jan;143(1):143-149. [doi: 10.1016/j.schres.2012.10.009] [Medline: 23146146]
48. Gleeson JF, Lederman R, Wadley G, Bendall S, McGorry PD, Alvarez-Jimenez M. Safety and privacy outcomes from a moderated online social therapy for young people with first-episode psychosis. Psychiatr Serv 2014 Apr 1;65(4):546-550. [doi: 10.1176/appi.ps.201300078] [Medline: 24687106]
49. Lederman R, Wadley G, Gleeson J, Bendall S, Alvarez-Jimenez M. Moderated online social therapy: designing and evaluating technology for mental health. ACM Trans Comput Hum Interact 2014 Feb 01;21(1):1-26. [doi: 10.1145/2513179]
50. Kuosmanen L, Välimäki M, Joffe G, Pitkänen A, Hätönen H, Patel A, et al. The effectiveness of technology-based patient education on self-reported deprivation of liberty among people with severe mental illness: a randomized controlled trial. Nord J Psychiatry 2009;63(5):383-389. [doi: 10.1080/08039480902825241] [Medline: 19308796]
51. Hätönen H, Suhonen R, Warro H, Pitkänen A, Välimäki M. Patients' perceptions of patient education on psychiatric inpatient wards: a qualitative study. J Psychiatr Ment Health Nurs 2010 May;17(4):335-341. [doi: 10.1111/j.1365-2850.2009.01530.x] [Medline: 20529184]
52. Kuosmanen L, Jakobsson T, Hyttinen J, Koivunen M, Välimäki M. Usability evaluation of a web-based patient information system for individuals with severe mental health problems. J Adv Nurs 2010 Dec;66(12):2701-2710. [doi: 10.1111/j.1365-2648.2010.05411.x] [Medline: 20722811]
53. Pitkänen A, Välimäki M, Kuosmanen L, Katajisto J, Koivunen M, Hätönen H, et al. Patient education methods to support quality of life and functional ability among patients with schizophrenia: a randomised clinical trial. Qual Life Res 2012 Mar;21(2):247-256. [doi: 10.1007/s11136-011-9944-1] [Medline: 21655934]
54. Anttila M, Välimäki M, Hätönen H, Luukkaala T, Kaila M. Use of web-based patient education sessions on psychiatric wards. Int J Med Inform 2012 Jun;81(6):424-433. [doi: 10.1016/j.ijmedinf.2012.02.004] [Medline: 22381805]
55. Laine A, Anttila M, Välimäki M. Modification of an internet-based patient education program for adults with schizophrenia spectrum disorder to suit adolescents with psychosis. Inform Health Soc Care 2016 Sep;41(3):230-246. [doi: 10.3109/17538157.2015.1008487] [Medline: 25710517]
56. Barnes E, Simpson S, Griffiths E, Hood K, Craddock N, Smith DJ. Developing an online psychoeducation package for bipolar disorder. J Ment Health 2011 Feb;20(1):21-31. [doi: 10.3109/09638237.2010.525565] [Medline: 21271823]
57. Smith DJ, Griffiths E, Poole R, di Florio A, Barnes E, Kelly MJ, et al. Beating Bipolar: exploratory trial of a novel Internet-based psychoeducational treatment for bipolar disorder. Bipolar Disord 2011;13(5-6):571-577. [doi: 10.1111/j.1399-5618.2011.00949.x] [Medline: 22017225]
58. Poole R, Simpson SA, Smith DJ. Internet-based psychoeducation for bipolar disorder: a qualitative analysis of feasibility, acceptability and impact. BMC Psychiatry 2012 Sep;12:139 [FREE Full text] [doi: 10.1186/1471-244X-12-139] [Medline: 22971042]
59. Barnes CW, Hadzi-Pavlovic D, Wilhelm K, Mitchell PB. A web-based preventive intervention program for bipolar disorder: outcome of a 12-months randomized controlled trial. J Affect Disord 2015 Mar 15;174:485-492. [doi: 10.1016/j.jad.2014.11.038] [Medline: 25554993]
60. Ben-Zeev D, Brenner CJ, Begale M, Duffecy J, Mohr DC, Mueser KT. Feasibility, acceptability, and preliminary efficacy of a smartphone intervention for schizophrenia. Schizophr Bull 2014 Nov;40(6):1244-1253. [doi: 10.1093/schbul/sbu033] [Medline: 24609454]
61. de Leeuw JRJ, van Splunteren P, Boerema I. Personal control in rehabilitation: An internet platform for patients with schizophrenia and their caregivers. Open J Psychiatry 2012;2(4):355-361. [doi: 10.4236/ojpsych.2012.224050]
62. Depp CA, Mausbach B, Granholm E, Cardenas V, Ben-Zeev D, Patterson TL, et al. Mobile interventions for severe mental illness: design and preliminary data from three approaches. J Nerv Ment Dis 2010 Oct;198(10):715-721 [FREE Full text] [doi: 10.1097/NMD.0b013e3181f49ea3] [Medline: 20921861]
63. Depp CA, Ceglowski J, Wang VC, Yaghouti F, Mausbach BT, Thompson WK, et al. Augmenting psychoeducation with a mobile intervention for bipolar disorder: a randomized controlled trial. J Affect Disord 2015 Mar 15;174:23-30. [doi: 10.1016/j.jad.2014.10.053] [Medline: 25479050]
64. Granholm E, Ben-Zeev D, Link PC, Bradshaw KR, Holden JL. Mobile Assessment and Treatment for Schizophrenia (MATS): a pilot trial of an interactive text-messaging intervention for medication adherence, socialization, and auditory hallucinations. Schizophr Bull 2012 May;38(3):414-425 [FREE Full text] [doi: 10.1093/schbul/sbr155] [Medline: 22080492]
65. Forchuk C, Donelle L, Ethridge P, Warner L. Client perceptions of the Mental Health Engagement Network: a secondary analysis of an intervention using smartphones and desktop devices for individuals experiencing mood or psychotic disorders in Canada. JMIR Ment Health 2015 Jan 21;2(1):e1. [doi: 10.2196/mental.3926]
66. Gottlieb JD, Romeo KH, Penn DL, Mueser KT, Chiko BP. Web-based cognitive-behavioral therapy for auditory hallucinations in persons with psychosis: a pilot study. Schizophr Res 2013 Apr;145(1-3):82-87. [doi: 10.1016/j.schres.2013.01.002] [Medline: 23410709]

67. Holländare F, Eriksson A, Lövgren L, Humble MB, Boersma K. Internet-based cognitive behavioral therapy for residual symptoms in bipolar disorder type II: A single-subject design pilot study. JMIR Res Protoc 2015;4(2):e44 [FREE Full text] [doi: 10.2196/resprot.3910] [Medline: 25908235]
68. Kane JM, Perlis RH, DiCarlo LA, Au-Yeung K, Duong J, Petrides G. First experience with a wireless system incorporating physiologic assessments and direct confirmation of digital tablet ingestions in ambulatory patients with schizophrenia or bipolar disorder. J Clin Psychiatry 2013 Jun;74(6):e533-e540. [doi: 10.4088/JCP.12m08222] [Medline: 23842023]
69. Latalova K, Prasko J, Kamaradova D, Jelenova D, Ociskova M, Sedlackova Z. Internet psychoeducation for bipolar affective disorder: basis for preparation and first experiences. Psychiatr Q 2014 Jun;85(2):241-255. [doi: 10.1007/s11126-013-9286-y] [Medline: 24307178]
70. Lauder S, Chester A, Castle D, Dodd S, Berk L, Klein B, et al. Development of an online intervention for bipolar disorder. www.moodswings.net.au. Psychol Health Med 2013;18(2):155-165. [doi: 10.1080/13548506.2012.689840] [Medline: 22712771]
71. Lauder S, Chester A, Castle D, Dodd S, Gliddon E, Berk L, et al. A randomized head to head trial of MoodSwings.net.au: an Internet based self-help program for bipolar disorder. J Affect Disord 2015 Jan 15;171:13-21. [doi: 10.1016/j.jad.2014.08.008] [Medline: 25282145]
72. Lieberman DZ, Swayze S, Goodwin FK. An automated Internet application to help patients with bipolar disorder track social rhythm stabilization. Psychiatr Serv 2011 Nov;62(11):1267-1269. [doi: 10.1176/appi.ps.62.11.1267] [Medline: 22211203]
73. Miklowitz DJ, Price J, Holmes EA, Rendell J, Bell S, Budge K, et al. Facilitated Integrated Mood Management for adults with bipolar disorder. Bipolar Disord 2012 Mar;14(2):185-197 [FREE Full text] [doi: 10.1111/j.1399-5618.2012.00998.x] [Medline: 22420594]
74. Murray G, Leitan ND, Berk M, Thomas N, Michalak E, Berk L, et al. Online mindfulness-based intervention for late-stage bipolar disorder: pilot evidence for feasibility and effectiveness. J Affect Disord 2015 Jun 1;178:46-51. [doi: 10.1016/j.jad.2015.02.024] [Medline: 25795535]
75. Nicholas J, Proudfoot J, Parker G, Gillis I, Burckhardt R, Manicavasagar V, et al. The ins and outs of an online bipolar education program: a study of program attrition. J Med Internet Res 2010;12(5):e57 [FREE Full text] [doi: 10.2196/jmir.1450] [Medline: 21169169]
76. Proudfoot J, Parker G, Manicavasagar V, Hadzi-Pavlovic D, Whitton A, Nicholas J, et al. Effects of adjunctive peer support on perceptions of illness control and understanding in an online psychoeducation program for bipolar disorder: a randomised controlled trial. J Affect Disord 2012 Dec 15;142(1-3):98-105. [doi: 10.1016/j.jad.2012.04.007] [Medline: 22858215]
77. Pijnenborg GHM, Withaar FK, Brouwer WH, Timmerman ME, van den Bosch J, Evans JJ. The efficacy of SMS text messages to compensate for the effects of cognitive impairments in schizophrenia. Br J Clin Psychol 2010 Jun;49(2):259-274. [doi: 10.1348/014466509X467828] [Medline: 19735607]
78. Rotondi AJ, Haas GL, Anderson CM, Newhill CE, Spring MB, Ganguli R, et al. A clinical trial to test the feasibility of a telehealth psychoeducational intervention for persons with schizophrenia and their families: intervention and 3-month findings. Rehabil Psychol 2005 Nov;50(4):325-336 [FREE Full text] [doi: 10.1037/0090-5550.50.4.325] [Medline: 26321774]
79. Rotondi AJ, Anderson CM, Haas GL, Eack SM, Spring MB, Ganguli R, et al. Web-based psychoeducational intervention for persons with schizophrenia and their supporters: one-year outcomes. Psychiatr Serv 2010 Nov;61(11):1099-1105. [doi: 10.1176/appi.ps.61.11.1099] [Medline: 21041348]
80. Rotondi AJ, Eack SM, Hanusa BH, Spring MB, Haas GL. Critical design elements of e-health applications for users with severe mental illness: singular focus, simple architecture, prominent contents, explicit navigation, and inclusive hyperlinks. Schizophr Bull 2015 Mar;41(2):440-448. [doi: 10.1093/schbul/sbt194] [Medline: 24375458]
81. Simon GE, Ludman EJ, Goodale LC, Dykstra DM, Stone E, Cutsogeorge D, et al. An online recovery plan program: can peer coaching increase participation? Psychiatr Serv 2011 Jun;62(6):666-669 [FREE Full text] [doi: 10.1176/appi.ps.62.6.666] [Medline: 21632737]
82. Todd NJ, Jones SH, Lobban FA. What do service users with bipolar disorder want from a web-based self-management intervention? A qualitative focus group study. Clin Psychol Psychother 2013;20(6):531-543. [doi: 10.1002/cpp.1804] [Medline: 22715161]
83. Todd NJ, Jones SH, Hart A, Lobban FA. A web-based self-management intervention for bipolar disorder 'living with bipolar': a feasibility randomised controlled trial. J Affect Disord 2014 Dec;169:21-29. [doi: 10.1016/j.jad.2014.07.027] [Medline: 25129531]
84. van der Krieke L, Emerencia AC, Aiello M, Sytema S. Usability evaluation of a web-based support system for people with a schizophrenia diagnosis. J Med Internet Res 2012;14(1):e24 [FREE Full text] [doi: 10.2196/jmir.1921] [Medline: 22311883]
85. van der Krieke L, Emerencia AC, Boonstra N, Wunderink L, de Jonge P, Sytema S. A web-based tool to support shared decision making for people with a psychotic disorder: randomized controlled trial and process evaluation. J Med Internet Res 2013;15(10):e216 [FREE Full text] [doi: 10.2196/jmir.2851] [Medline: 24100091]
86. Wenze SJ, Armey MF, Miller IW. Feasibility and acceptability of a mobile intervention to improve treatment adherence in bipolar disorder: A pilot study. Behav Modif 2014 Jan 8;38(4):497-515. [doi: 10.1177/0145445513518421] [Medline: 24402464]

87. Pew Research Center. Communications Technology in Emerging and Developing Nations. Washington, DC: Pew Research Center; 2015 Mar 19. URL: http://www.pewglobal.org/2015/03/19/1-communications-technology-in-emerging-and-developing-nations/ [accessed 2015-09-18] [WebCite Cache ID 6bdOyiYhn]
88. Brian RM, Ben-Zeev D. Mobile health (mHealth) for mental health in Asia: objectives, strategies, and limitations. Asian J Psychiatr 2014 Aug;10:96-100. [doi: 10.1016/j.ajp.2014.04.006] [Medline: 25042960]
89. Farrington C, Aristidou A, Ruggeri K. mHealth and global mental health: still waiting for the mH2 wedding? Global Health 2014;10:17 [FREE Full text] [doi: 10.1186/1744-8603-10-17] [Medline: 24670011]
90. Bucci S, Barrowclough C, Ainsworth J, Morris R, Berry K, Machin M, et al. Using mobile technology to deliver a cognitive behaviour therapy-informed intervention in early psychosis (Actissist): study protocol for a randomised controlled trial. Trials 2015;16(1):404 [FREE Full text] [doi: 10.1186/s13063-015-0943-3] [Medline: 26357943]
91. Lobban F, Dodd AL, Dagnan D, Diggle PJ, Griffiths M, Hollingsworth B, et al. Feasibility and acceptability of web-based enhanced relapse prevention for bipolar disorder (ERPonline): trial protocol. Contemp Clin Trials 2015 Mar;41:100-109 [FREE Full text] [doi: 10.1016/j.cct.2015.01.004] [Medline: 25602581]
92. Hidalgo-Mazzei D, Mateu A, Reinares M, Undurraga J, Bonnín CM, Sánchez-Moreno J, et al. Self-monitoring and psychoeducation in bipolar patients with a smart-phone application (SIMPLe) project: design, development and studies protocols. BMC Psychiatry 2015;15:52 [FREE Full text] [doi: 10.1186/s12888-015-0437-6] [Medline: 25884824]

Abbreviations

IAPT: Improving Access to Psychological Therapies
NHS: National Health Service
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
SMI: severe mental health problems

Edited by G Eysenbach; submitted 19.10.15; peer-reviewed by N Thomas, J Torous, A Rotondi; comments to author 14.02.16; revised version received 15.03.16; accepted 03.04.16; published 31.05.16.

Please cite as:
Berry N, Lobban F, Emsley R, Bucci S
Acceptability of Interventions Delivered Online and Through Mobile Phones for People Who Experience Severe Mental Health Problems: A Systematic Review
J Med Internet Res 2016;18(5):e121
URL: http://www.jmir.org/2016/5/e121/
doi:10.2196/jmir.5250
PMID:27245693

Original Paper

Acceptability, Feasibility, and Cost of Telemedicine for Nonacute Headaches: A Randomized Study Comparing Video and Traditional Consultations

Kai Ivar Müller[1,2]; Karl Bjørnar Alstadhaug[2,3]; Svein Ivar Bekkelund[1,2]

[1]Department of Neurology, University Hospital of North Norway, Tromsø, Norway

[2]Department of Clinical Medicine, UiT – The Arctic University of Norway, Tromsø, Norway

[3]Department of Neurology, Nordland Hospital Trust, Bodø, Norway

Corresponding Author:
Kai Ivar Müller
Department of Neurology
University Hospital of North Norway
Sykehusvegen 38
Tromsø, N-9038
Norway
Phone: 47 77627073
Fax: 47 77627074
Email: kai.ivar.muller@unn.no

Abstract

Background: The feasibility of telemedicine in diagnosing and treating nonacute headaches, such as primary headaches (migraine and tension-type) and medication-overuse headaches has not been previously investigated. By eliminating the need of travel to specialists, telemedicine may offer significant time and money savings.

Objectives: Our objective was to estimate the acceptance of telemedicine and investigate the feasibility and cost savings of telemedicine consultations in diagnosing and treating nonacute headaches.

Methods: From September 2012 to March 2015, nonacute headache patients from Northern Norway who were referred to neurologists through an electronic application system were consecutively screened and randomized to participate in either telemedicine or traditional specialist visits. All patients were consulted by two neurologists at the neurological department in Tromsø University Hospital. Feasibility outcomes were compared between telemedicine and traditional groups. Baseline characteristics and costs were then compared between rural and urban patients. Travel costs were calculated by using the probabilistic method of the Norwegian traveling agency: the cheapest means of public transport for each study participant. Loss of pay was calculated based on the Norwegian full-time employee's average salary: < 3.5 hours=a half day's salary, > 3.5 hours spent on travel and consultation=one day's salary. Distance and time spent on travel were estimated by using Google Maps.

Results: Of 557 headache patients screened, 479 were found eligible and 402 accepted telemedicine participation (83.9%, 402/479) and were included in the final analyses. Of these, 202 received traditional specialist consultations and 200 received telemedicine. All patients in the telemedicine group were satisfied with the video quality, and 198 (99%, 198/200) were satisfied with the sound quality. The baseline characteristics as well as headache diagnostics and follow-up appointments, and the investigation, advice, and prescription practices were not statistically different between the two randomized groups. In addition, telemedicine consultations were shorter than traditional visits (38.8 vs 43.7 min, $P<.001$). The travel cost per rural individual (292/402, 73%) was €249, and estimated lost income was €234 per visit. The travel cost in the urban area (110/402, 27%) was €6, and estimated lost income was €117 per visit. The median traveling distance for rural patients was 526 km (range 1892 km), and the median traveling time was 7.8 hours (range 27.3 hours). Rural patients had a longer waiting time than urban patients (64 vs 47 days, $P=.001$), and fewer women were referred from rural areas ($P=.04$). Rural women reported higher pain scores than urban women ($P=.005$).

Conclusion: Our study shows that telemedicine is an accepted, feasible, time-saving, and cost-saving alternative to traditional specialist consultations for nonacute headaches.

Trial Registration: Clinicaltrials.gov NCT02270177; http://clinicaltrials.gov/ct2/show/NCT02270177 (Archived by WebCite at http://www.webcitation.org/6hmoHGo9Q)

(J Med Internet Res 2016;18(5):e140) doi:10.2196/jmir.5221

KEYWORDS

headache; management; consultation; telemedicine; burden; cost; feasibility; rural; randomization

Introduction

Nonacute headaches are among the most frequent disorders in humans [1,2]. The global burden of these headaches (eg, migraine, tension-type, and medication overuse headaches (MOH)) account for the most common neurological cause of disability-adjusted life years and years lived with disability (45.1% and 69.3% of the total in 2013, respectively) [2,3]. Moreover, nonacute headaches represent a frequent cause of referrals to neurologic outpatient clinics and are considered to be a cause of the third largest neurological health cost in Europe, estimated annually at €43,514 million in 2010 [4,5].

Patients with nonacute conditions need a referral to be accepted to a specialist consultation according to Norwegian health laws. Nonacute headaches usually occur for at least four weeks without any clinical or radiological signs of structural intracranial pathology [6]. Patients in our study are diagnosed according to the second version of the International Classification of Headache Disorders 2 (ICHD-2) [7]. The most frequent primary headaches are migraines and tension-type headaches, while an important cause of chronic headache is MOHs. These also constitute the majority of nonacute headaches.

Northern Norway's physical geography is extensive with many sparsely populated areas. Access to specialist health care becomes cumbersome, expensive, and time consuming for many patients because of variable weather conditions. Efforts to facilitate easier access to specialists for headache patients are furthermore obstructed by tight health budgets. Accordingly, the rules and regulations surrounding the practice of telemedicine in Norway are addressed in a government circular letter of 2001 [8]. In general, this document states that consultations, diagnostics, treatments, and safety issues with telemedicine are governed by the same principles as traditional face-to-face consultations.

Telemedicine may help reduce the burden for neurological patients who live in rural or underserved areas [9,10], and it is considered equal in quality to traditional visits among different medical professions [11]. Guidelines for telemedicine in a second opinion for headaches have been suggested [12], but the feasibility of telemedicine being used as a tool for outpatient headache specialist consultations has only been addressed in case series [13-16].

A 2013 review of telemedicine interventions for somatic diseases found that 23% of papers (7/31) documented effectiveness or cost-effectiveness and 42% (13/31) showed promising results [17]. However, the evidence for health costs and patient acceptability in telemedicine was nonconclusive according to a 2013 Cochrane review [16]. Lack of such evidence hampers implementation of new information and communication technologies at the expense of patients' needs [18].

The aim of this study was to evaluate patient acceptance. We also sought to investigate the feasibility of using telemedicine in order to hinder access barriers by being independent of the patient's and headache specialist's location. Additionally, our study also estimated cost savings by consulting headache patients via telemedicine in the spectrum of headaches referred from general practice.

Methods

Study Design

Feasibility and economic evaluations are part of an ongoing open-labeled noninferiority randomized clinical trial (Clinicaltrial.gov id. NCT02270177). In this trial, specialist telemedicine visits versus traditional specialist visits for headache sufferers were compared.

Study Population

The study was conducted in the neurological outpatient clinic at the University Hospital of Northern Norway in Tromsø city. This facility serves 190,726 inhabitants in Troms county and upper Nordland county, which are distributed over an area of 25,877 km^2 [19]. People living in areas further toward north, as in Finnmark and Svalbard, are also served by the same clinic. Finnmark has 75,605 inhabitants spread over a land area of 48,618 km^2, and Svalbard has a Norwegian population of 2180 inhabitants spread over an area of 61,022 km^2 [19]. In general, individuals from Finnmark and Svalbard travel by plane. Individuals from Troms and Nordland travel by car, bus, or boat.

We screened all patients who were referred to specialists for headaches from September 30, 2012 to March 30, 2015. The inclusion criteria were as follows:

1. Females and males aged ≥16 and ≤65 years

2. Referred to a neurologist for headache diagnostic clarification and/or treatment

3. Lack of abnormal findings on either clinical neurological examination, reported by the referring doctor, or by imaging of the brain suggestive of a secondary cause

4. Waiting time ≤4 months from date of the referral letter

5. Speaking Norwegian language.

To prevent working with patients already diagnosed with headaches, those who had been evaluated by a neurologist for 2 years before referral were excluded.

The legal age of consent in Norway is 16 years. To reduce the risk of including patients with secondary headaches and to recruit from a working population, patients above 65 years of age were excluded. In general, eligibility criteria were set to prevent working with participants unsuitable for telemedicine. Therefore, referrals of acute and secondary headaches, or

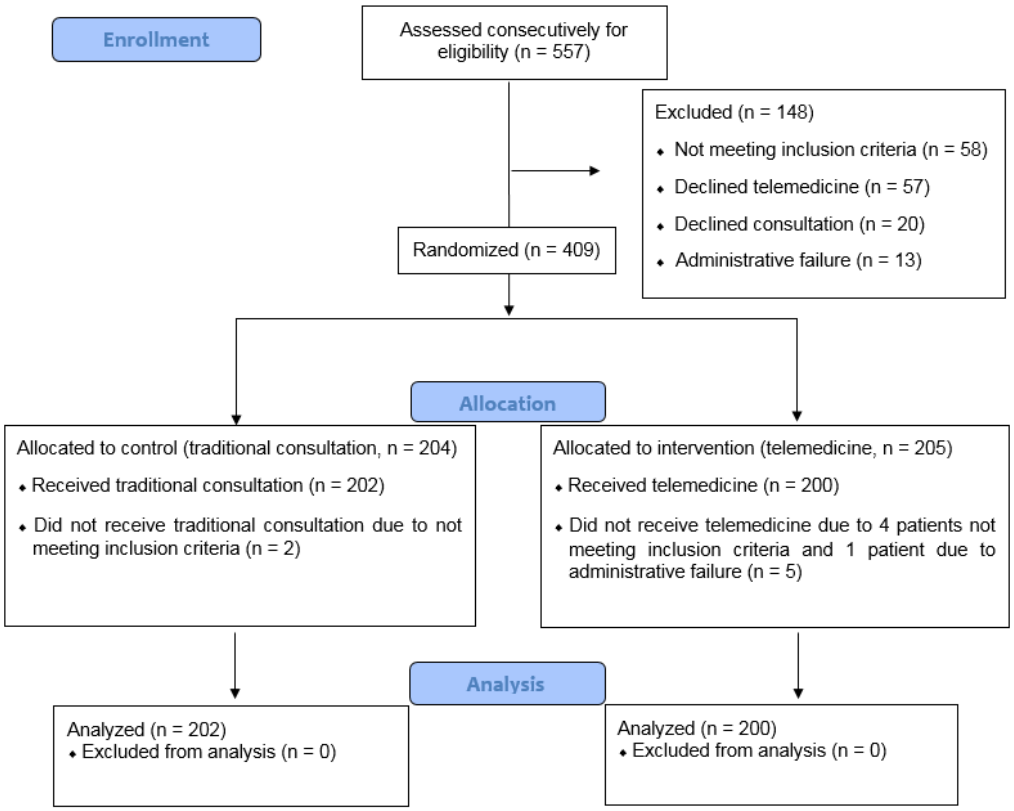
evidence of such, were excluded. For recruitment purposes and to prevent outdated information in the referral letters, the waiting time for a consultation was set to 4 months or less. Speaking the Norwegian language was considered a necessity to ensure proper and accurate communication.

Recruitment and Randomization

All eligible patients were identified through an electronic application system present in the distributed information system and patient system for hospitals from DIPS AS (DIPS) at the hospital. These individuals received an information letter and were interviewed by a study coordinator for eligibility criteria, telemedicine acceptance, and study participation. Individuals who did not meet the criteria or did not want to participate were transferred back to the traditional consultation setup. Next, candidates were called for final information and participation. Volunteers then received a consent form, a questionnaire, and a summoning letter. A study coordinator consecutively scheduled the participants for consultation, and they all met at the neurological outpatient department at Tromsø University Hospital.

A study nurse subsequently called the randomization office at the hospital and followed each patient to the consultation. The consultations took place between 9 am and 3 pm. Patients were block randomized by using Microsoft Access. The block sizes varied among four, six, and eight, and stratification was made on each neurologist. Furthermore, patient preparation time and consultation time were recorded. A study flow chart is given in Figure 1.

Figure 1. Flow of participants through the study.



Equipment

Video consultations were performed by using a videoconference system (Cisco C40 integrator package, Cisco C40 Integrator Multisite, Cisco Precision HD 1080p 12xcamera, NEC X551s 55” LED monitor, Audio-Technica ceiling microphones and JBL LSR2325P active speakers, Integrator Package C40 dual display option, and Cisco touch-control device for C Series) that was installed in one office. The neurologist consulted the patients from two other offices via a Cisco EX60 unit with an InTouch panel. Moreover, face-to-face consultations were performed in a traditional manner from the same offices.

Outcome Variables and Measurements

All data, including background variables, were recorded by structured interviews using an administrative protocol. Nonacute headaches were categorized according to the diagnosis and frequency into three groups: <7 days, 7-15, and ≥15 days per month within the last 3 months before consultation. Headaches occurring for ≥15 days per month for more than 3 months were defined as chronic while those occurring for <7 or 7-14 days were termed episodic [7].

A horizontal visual analog scale (VAS) ranging from 0 to 10, which has demonstrated validity and reliability for many pain conditions including headaches, has been in use both in research and clinical settings [20]. This method was used to measure headache intensity (0=no pain and 10=worst possible pain). Measurements were performed by the headache specialist during consultation, where a six-item headache impact test (HIT-6) with possible scores from 36-78 was used to assess the impact of headaches on daily life of the patients. In general, a score

above 50 is considered high. HIT-6 is validated and reliable for assessing headache impact with a Cronbach alpha of.89-.90 [21]. In our study, the six items associated with HIT-6 scores had good reliabilities (all Cronbach alpha=.84). Two experienced neurologists (Kai Ivar Müller, Svein Ivar Bekkelund) conducted all the consultations. Meanwhile, the telemedicine participants were guided out of sight to the consultant.

To evaluate the feasibility (a-i), cost (j-k), and travel savings (l) of telemedicine in urban and rural patients, the following outcome variables were selected: (a) eligible patients' acceptance of telemedicine, (b) dropout rates and causes of dropouts (medical or technical reasons), (c) participants' satisfaction with video quality and sound quality (patients were asked if they were "satisfied" with the video quality and sound quality after each telemedicine consultation; "Yes" or "No, why not?"), (d) technical errors with the telemedicine equipment, and (e) specialist consultation time in minutes. Descriptive variables included (f) diagnostic investigations (CT, MRI, and other), (g) nonpharmacologic advices, (h) number of prescriptions, and (i) follow-up appointments. Cost and travel variables included (j) cost of travel in euros (€) based on the Norwegian Patient Travel Agency probabilistic method [22], (k) estimated loss of income (<3.5 hours=a half day's salary, >3.5 hours spent on travel and consultation=one day's salary), (l) traveling distance in kilometers and traveling time in hours as estimated by Google Maps.

The Norwegian Patient Travel Agency calculated the cheapest means of public transport to and from Tromsø University Hospital for every patient. We collected the cheapest cost from the Norwegian Patient Travel Agency from every participating patient in every municipality of our area. Earnings were calculated based on the Norwegian full-time employee's average salary, which was €4.681 per month in 2014 [23]. We adjusted all costs to the consumer price index (CPI) January 1, 2015 from Statistics Norway and converted Norwegian kroner into euros by using the exchange rate of one Norwegian krone per euro from the Norwegian Bank on December 31, 2014. Since the telemedicine equipment is in use for many different purposes at the University Hospital in Tromsø, we did not include the cost of the equipment, its installation, or its maintenance.

Statistical Analyses

The computer program SPSS version 21 was used to analyze the data. Continuous variables were tested for normal distribution with the Shapiro-Wilk test, skewness, and kurtosis. Normally distributed variables were presented as mean (SD) and categorical data as numbers and percentages. Independent sample t-tests were used to compare continuous variables and chi-square tests were used to compare categorical variables between groups. Yates' continuity correction was used for 2 × 2 tables. In addition, skewed variables are presented as median and range, while the Mann-Whitney U test was used to compare continuous variables. All variables were first tested in univariate models. Statistical significance was defined as $P<.05$.

Hierarchical multiple linear regression was used to assess the ability to predict VAS pain scores of patient's from rural locations when adjusted for other variables. After sex and age, only variables that could be associated with changes in pain scores and that showed significance in the univariate analysis were selected. Hierarchical multiple linear regression was also used to assess the ability to predict the waiting time of patients from rural locations. We adjusted for age, sex, and other predictors. Furthermore, nonparametric variable waiting time was log transformed in the regression model. Normal probability plot (P-P) of the regression standardized residuals, histogram, and scatterplots were used to assure normality, linearity, and residual independence as well as to eliminate outliers in both models. In addition, multicollinearity was checked by tolerance and variance inflation factor (VIF) in both models.

Consent and Ethical Approval

Oral and written consent from all participants was obtained. The participants' mental and physical integrity has been fully respected and safeguarded in accordance with the Helsinki Declaration [24]. The study was approved by the Norwegian National Committee for Medical and Health Research Ethics (REC), number 2009/1430/REK.

Results

All referred nonacute headache patients (N=557) were screened, and 402 of 486 (82.7%) candidates were included in the final analyses (Figure 1). From 409 specialist consultations, 7 (1.7%, 7/409) patients were excluded. Of the eligible individuals, 402/479 (83.9%) accepted telemedicine, 57/479 (11.9%) did not approve telemedicine, and 20/479 (4.2%) declined specialist care. One patient in the traditional consultation group was excluded due to a transient left-sided hemiparesis observed in his headache history. Second patient was excluded as the maximum waiting time was exceeded. Also, in the telemedicine group, one patient was excluded due to a cystic cerebellar lesion seen on an MRI. Another had ataxia, and the third had suspect subcutaneous tumors on his scalp. Additionally, one patient had been seen by a neurologist within a 2-year period prior to the referral, and another patient was excluded because the consulting telemedicine room was occupied.

All included patients underwent only MRI/CT brain scan (205/402 participants, 51.0%), only neurological examination (46/402, 11.4%), or both (151/402, 37.6%). The study participants did not differ significantly from the excluded individuals by sex and age (n=155, P=.17 and .41, respectively).

Migraines (219/402 patients, 54.5%) and probable MOH (73/402 patients, 18.2%) were the most common referral reasons (Figure 2). The most frequent comorbid disorders were chronic neck pain (188/402, 46.8%) and insomnia (126/402, 31.3%), followed by hypertension (36/402, 9.0%), asthma (21/402, 5.2%), depression (18/402, 4.5%), and hypothyreosis (15/402, 3.7%). Two participants had a history of cerebrovascular disease, and one had epilepsy. Patients who underwent traditional consultations did not significantly differ statistically by demographics and clinical characteristics as compared to the patients in the telemedicine group (Table 1).

Table 1. Baseline characteristics of patients with headache who were referred to specialists.

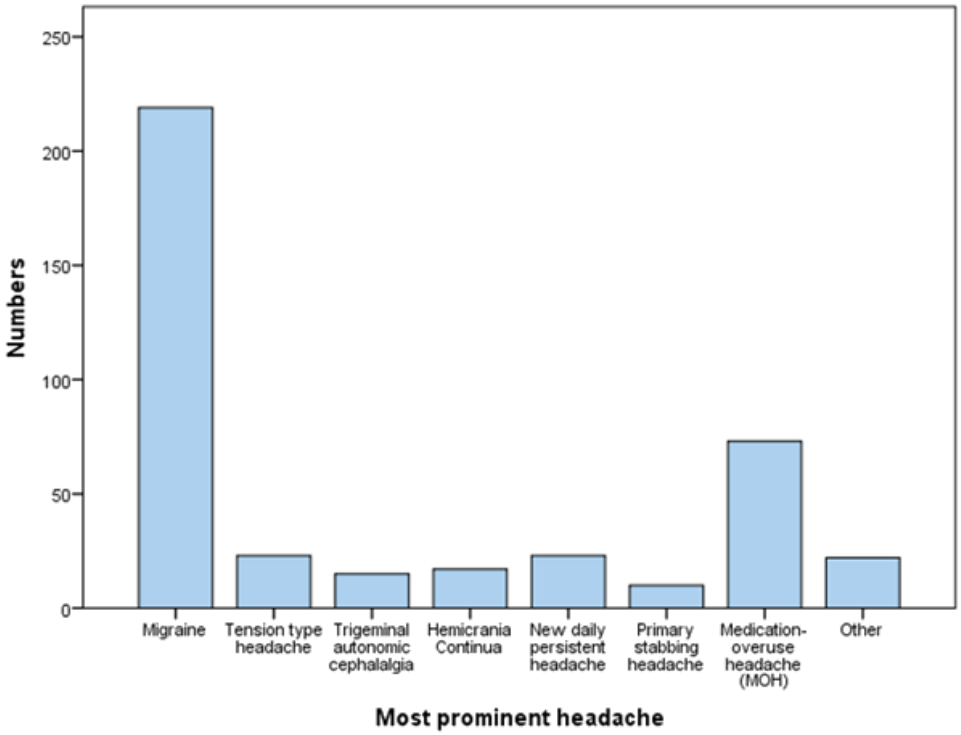
Consultation form Characteristic		TelemedicineN=200	Traditional N=202	*P*-value
Females, n (%)		148 (74.0)	153 (75.7)	.77
Married/cohabitating, n (%)		117 (58.5)	135 (67.0)	.09
Age, years, mean (SD)		36.0 (13.0)	38.0 (13.7)	.12
Education, years (SD)		13.5 (3.0)	13.8 (3.1)	.22
Employment full time, n (%)		128 (64.0)	125 (61.9)	.74
Unemployed, n (%)		11 (5.5)	19 (9.4)	.19
Sick leave due to headache, n (%)		58 (29.0)	62 (30.7)	.79
[a] **Headache days/month, n (%)**				.51
	> 15 days	120	113	
	7-15 days	41	40	
	< 7 days	39	49	
HIT-6, mean (SD)		64.1 (6.1)	64.0 (6.1)	.49
VAS, mean (SD)		7.1 (2.2)	6.9 (2.1)	.82
Most prominent headache				
	Migraine, n (%)	106 (53.0)	113 (55.9)	.62
	TTH, n (%)	15 (7.5)	8 (4.0)	.19
	MOH, n (%)	35 (17.5)	38 (18.8)	.83
Waiting time, days, median (range)		63 (117)	59 (119)	.34
Travel cost per patient €, median (range)		83 (437)	83 (437)	.96
[b]Lost pay per patient, €, median (range)		234 (117)	234 (117)	.87

SD: standard deviation; HIT-6: headache impact test-6; VAS: visual analog scale; TTH: tension type headache; MOH: Medication overuse headache.

[a]Last 3 months before headache consultation.

[b]Corresponding to a half day's salary in the urban area and a day's salary in rural areas.

Figure 2. Overview of the most prominent headaches, n=402.



Feasibility Variables

The feasibility variables of the randomized groups are shown in Table 2. Telemedicine had shorter consultations (*P*<.001), especially for men (35.3 vs 43.0 min, *P*=.001). Men had shorter telemedicine consultations than women (35.3 vs 40.0, *P*=.002). Meanwhile, there were no considerable differences in consultation time between men and women in the traditional group (43.0 vs 44.0, *P*=.63). Other feasibility parameters did not differ between the groups (*P*>.05; Table 2). All participants were satisfied with the video quality, and 198/200 (99.0%) were satisfied with the sound quality. In the first case, there was no sound transmission, and the patient had to be consulted through video and the loudspeaker on the ordinary telephone at the office. In the second case, the patient felt the sound was too loud, and the study nurse was called upon to decrease the volume.

A significantly lower share of women were referred from rural areas compared with the urban area (*P*=.04), and the rural women had significantly higher pain scores (*P*=.005). A hierarchical multiple regression model of VAS predictors showed that sex and probable MOH explained 3.5% of the VAS variance (*P*=.003) while the whole model explained 5.1% (*P*=.01). For each rural patient, there was a 0.6 increase on the VAS scale (*P*=.01) when adjusted for age, sex, and diagnosis (Table 3). We did not find other significant VAS predictors than sex, probable MOH, and patient location.

Table 2. Clinical and feasibility evaluation of telemedicine consultations in patients who were referred to specialists for headaches.

Consultation form		Telemedicine, N=200	Traditional, N=202	*P*-value
Clinical variables				
	Change in diagnosis, n (%)	43 (21.5)	40 (19.8)	.77
	Additional diagnosis, n (%)	109 (54.5)	118 (58.4)	.49
	Additional MRI/CT, n (%)	74 (37.0)	70 (34.7)	.70
	Nonpharmacologic advice, n (%)	162 (81.0)	170 (84.6)	.41
	Prescriptions, n (%)	164 (82.0)	166 (82.2)	1.00
	Follow-up by GP, n (%)	136 (68.0)	129 (63.9)	.44
	Follow-up by neurologist, n (%)	6 (3.0)	6 (3.0)	1.00
Outcome variables				
	Dropout, medical reasons, n (%)	2 (1.0)	1 (0.5)	NA
	Dropout, technical failure, n (%)	1 (0.5)	0	NA
	Minor technical issues, n (%)	21 (10.5)	NA	NA
	Satisfied with sound quality, n (%)	198 (99.0)	NA	NA
	Satisfied with video quality, n (%)	200 (100.0)	NA	NA
	Consultation time, min, mean (SD)	38.8 (9.5)	43.7 (12.3)	< .001
	Males, min (SD)	35.3 (8.3)	43 (13.1)	.001
	Females, min (SD)	40.0 (9.6)	44.0 (12.0)	.002
	[a]Preparation to visit, min (range)	14.0 (44)	14.0 (29)	.56

GP: general practitioner; NA: not applicable; SD: standard deviation

[a]The time used by the nurse before each consultation.

Table 3. Hierarchical linear regression model to assess rural location as a predictor of VAS when controlled for confounders.

Step		[a]B (standard error)	95% CI	*P*-value
Step 1				
	Constant	5.786 (.578)	4.649 to 6.922	<.001
	Age	-.003 (.008)	-0.19 to .013	.71
	Sex	.699 (.250)	.207 to 1.191	.01
	MOH	.649 (.279)	.101 to 1.198	.02
Step 2				
	Constant	5.220 (.614)	4.013 to 6.427	<.001
	Age	-.003 (.614)	-.019 to .013	.69
	Sex	.770 (.250)	.278 to 1.261	.002
	MOH	.637 (.277)	.092 to 1.182	.02
	Urban/rural	.625 (.240)	.153 to 1.096	.01

CI: confidence intervals; MOH: Probable medication overuse headache.

[a]Unstandardized coefficients.

Patients in the rural group had significantly longer waiting times for specialist consultations (median=64 days, range=120) than patients in the urban group (median=47 days, range=112, *P*=.001). A hierarchical multiple linear regression model to assess rural location and predict waiting time after adjusting parameters for age and sex showed a total variance of 4.4% as explained by the model (*P*=.001). Age, sex, marital status, education, employment, sick leave, HIT-6, migraine, tension type headache, and probable MOH were insignificant predictors of waiting time.

Cost and Travel Estimates

A comparison between urban and rural participants is given in Table 4. The potential of lowest estimated travel cost, loss of pay, and travel burden compared between urban and rural patients is summarized in Table 5. Travel cost per patient was

notably higher in rural areas (outside Tromsø city; median=€249.0, range=409) as compared with urban areas (in Tromsø city; €6.0 per patient, *P*<.001).

Table 4. Characteristics of headache patients from urban and rural areas who were referred to specialists.

Characteristic		[a]Urban, N=110	[b]Rural, N=292	*P*-value
Females, n (%)		91 (82.7)	210 (71.9)	.04
Married/cohabitating, n (%)		69 (62.7)	183 (72.6)	.31
Age, years, mean (SD)		36.3 (13.0)	37.2 (13.6)	.54
Education, years (SD)		14.0 (3.0)	13.6 (3.0)	.24
Employment full time, n (%)		70 (63.6)	183 (62.7)	.95
Sick leave due to headache, n (%)		28 (25.5)	92 (31.5)	.33
[c] **Headache days/month, n (%)**				
	> 15 days	62 (56.4)	171 (58.6)	.11
	7-15 days	29 (26.4)	52 (17.8)	NA
	< 7 days	19 (17.3)	69 (23.6)	NA
[d] **HIT-6, mean (SD)**		63.3 (7.1)	64.4 (5.7)	.15
	Female HIT-6	63.9 (6.8)	64.7 (5.5)	.80
	Male HIT-6	60.3 (7.9)	63.6 (6.0)	.047
[d] **VAS, mean (SD)**		6.6 (2.4)	7.2 (2.1)	.04
	Female VAS	6.6 (2.4)	7.5 (1.9)	.005
	Male VAS	6.6 (2.3)	6.4 (2.4)	.82
Waiting time, days, median (range)		47 (112)	64 (120)	.001

[a]Urban: patients from Tromsø city.

[b]Rural: patients living outside Tromsø city.

[c]Last 3 months before headache consultation.

[d]HIT-6: headache impact test-6; SD: standard deviation; VAS: visual analog scale.

Table 5. Cost and travel burden for headache patients from urban and rural areas who were referred to specialists.

Cost	[a]Urban, N=110	[b]Rural, N=292	*P*-value
Travel cost per patient, €, median (range)	[c]6	249 (409)	<.001
[d]Loss of earnings per patient, euro	117	234	<.001
Total cost per patient, euro	123	483	<.001
Travel, kilometers, median (range)	12 (69)	526 (1892)	<.001
Travel time, hours, median (range)	0.4 (1.2)	7.8 (27.3)	<.001

[a]Urban: patients from Tromsø city.

[b]Rural: patients living outside Tromsø city.

[c]The price of a bus ticket in Tromsø city.

[d]Corresponding to a half day's salary in the urban area and a day's salary in rural areas.

Discussion

Headache Patients' Acceptability of Telemedicine

There was high level of acceptance for telemedicine, and only 1% (2/200) was unsatisfied with the technical quality of the consultation in our study. These findings concur with a 2009/2010 survey from a remote area, including 1816 individuals, which reports a social acceptance of telehealth at 77.7% (1356/1745) and a confidence level at 65.8% (1146/1742) [25]. On the other hand, another survey conducted in 2008 with a total of 1634 individuals spread over different European countries shows that 42%-81% (964/1634) of patients with headache are unsatisfied with their health care [9]. Additionally, a similar European survey from 2013 shows that 48 % (929/1935) are unsatisfied with the headache management [9]. The authors concluded that access and availability to health care had previously not been given enough attention when analyzing

the burden of headaches [9]. Despite having maintained a strict maximum waiting time of 4 months in our study, patients from rural areas waited significantly longer for headache consultations (Table 4). The potential burden of long and cumbersome traveling together with difficult journey logistics may explain this. In these areas, telemedicine may become a countervailing technology that can remodel access and availability barriers for headaches.

Telemedicine Feasibility for Nonacute Headaches

Despite the importance of investigating the feasibility of new technology, only case series exist regarding the use of telemedicine for new headache referrals [13,14]. Previous studies had mainly focused on education programs and follow-up treatments [26-28]. We found that the duration of the telemedicine consultations was 5 min (11%, 38.8 vs 43.7 min) shorter than that of traditional consultations. However, there were no statistical differences between telemedicine visits and traditional visits with respect to headache diagnoses, investigations, advices, prescriptions, and follow-up appointments. Telemedicine caused few dropouts. Minor technical errors of sound and video transmissions were quickly dealt with and did not influence the consultations or participants' satisfaction (Table 2). Almost all patients randomized to telemedicine completed the study without medical or technical problems. Due to no statistical differences of feasibility in favor of traditional consultations, reduced consultation time may also economically favor telemedicine.

Moreover, most women have more home responsibilities than men. Reduced consultation time along with less traveling may therefore also economically favor the patients' families, especially in rural areas. Additionally, we found that men had telemedicine consultations that were almost 5 min shorter than those of women. This finding may be attributed to the fact that men have a tendency to consult less than women [29]. Nonetheless, we did not find gender differences in the traditional group. Safety and quality of shorter consultations should be evaluated in prospective follow-up studies.

The number of women referred to specialists from rural areas being lower than those from urban areas and rural women having been referred seemingly having higher headache burdens may indicate that rural women consult general practitioners (GPs) for headaches less often than men. This phenomenon may be a consequence of long traveling logistics to see a headache specialist that interferes with women's working routines and childcare and home responsibilities. However, headache intensity remained significant even after adjustment for sex, employment, education, and age. Different attitudes and referral practices between rural and urban GPs is another explanation, but this idea needs further investigation. It is well known that headaches are underdiagnosed, often misdiagnosed, and suboptimally treated [30-32].

Cost Savings of Telemedicine Headache Consultations

Among headache patients referred to specialists in Northern Norway, almost three-fourths of them live in rural areas. For each of these patients with headache, our analyses showed that telemedicine consultations may result in a median travel reduction of €249 as well as a saving of €234, corresponding to one day's work (Table 5).

As we calculated the lowest possible journey expenses per patient, the amount of actual travel costs may be higher. Previous reviews of literature revealed diverging evidence on the cost-effectiveness and patient acceptability of telemedicine [16,33,34]. A cost-consequence study of 2094 patients in a randomized controlled trial found that joint teleconsultations referred by GPs to orthopedists, urologists, ear-nose-throat specialists, gastroenterologists, and other hospital specialists were more expensive by €138 than traditional visits [35]. Even then, the study concluded that total costs for patient travel and loss of pay would be decreased. Meanwhile, in a randomized controlled trial published in 2001, Chua et al found that telemedicine was not cost-effective for neurological outpatient consultations, but none of the patients had headaches [36]. As expected, our study demonstrated that travel costs are highest in rural areas, especially in those without specialist coverage. In Norway, a minimum of traveling expenses is paid by the patients themselves, which may be tax-deductible, but the regional health authority covers most of it and is also responsible for hospital budgets in the region. Reduced travel costs may allow allocation of more money for treatment.

Advantages and Limitations

A relatively large number of participants, a high response rate, randomization without statistically significant variation in demographics, headache characteristics and the burden between telemedicine and traditional groups (Table 1), and the real-life recruitment design are the main advantages of the this study. Although, organizing teleconsultations inside the hospital provides standardized conditions and promotes internal validity, it is less realistic than local evaluation of the patients. The fact that patients participated in a study, having increased confidence in the treating specialists, may have biased the results. Moreover, the lack of a placebo control group and blinding should not be underestimated; however, designing such a study would be difficult.

Conclusions

Our study supports telemedicine as a socially accepted, feasible, and cost-saving technology for diagnosing and treating nonacute headaches. A modern interdependent health care system providing simpler, faster, and cheaper services for patients with headache could be implemented. We are obligated to organize and conduct health care systems efficiently, based more on patients' needs than on the traditional paradigms. However, the quality and safety of employing telemedicine for patients with headache should be assessed in prospective follow-up studies.

Acknowledgments

We thank Jorun Willumsen, Anna-Kirsti Kvitnes, Marlen Lauritsen, Nora Bekkelund, Thomas Krogh, Karin Flatekval Eines, Torill Erdahl, Marianne Røst, and Grethe Berg Johnsen for their skillful help. This study was supported by grants from the Northern Norway Regional Health Authority (Helse Nord RHF).

Conflicts of Interest

None declared.

References

1. Rasmussen BK, Jensen R, Schroll M, Olesen J. Epidemiology of headache in a general population--a prevalence study. J Clin Epidemiol 1991;44(11):1147-1157. [Medline: 1941010]
2. Global Burden of Disease Study 2013 Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990-2013: a systematic analysis for the Global Burden of Disease Study 2013. Lancet 2015 Aug 22;386(9995):743-800 [FREE Full text] [doi: 10.1016/S0140-6736(15)60692-4] [Medline: 26063472]
3. Institute for Health Metrics and Evaluation (IHME). URL: http://ghdx.healthdata.org/global-burden-disease-study-2013-gbd-2013-data-downloads [accessed 2016-01-22] [WebCite Cache ID 6ein91Avj]
4. Bekkelund SI, Albretsen C. Evaluation of referrals from general practice to a neurological department. Fam Pract 2002 Jun;19(3):297-299 [FREE Full text] [Medline: 11978723]
5. Olesen J, Gustavsson A, Svensson M, Wittchen H, Jönsson B, group Cs. The economic cost of brain disorders in Europe. Eur J Neurol 2012 Jan;19(1):155-162. [doi: 10.1111/j.1468-1331.2011.03590.x] [Medline: 22175760]
6. Frishberg B, Rosenberg J, Matchar D, McCrory D, Pietrzak M, Rozen T, et al. Practice Guidelines. Minneapolis: American Academy of Neurology; 2000. Evidence-based guidelines in the primary care setting: neuroimaging in patients with nonacute headache URL: http://protocols.xray.ufl.edu/live_protocols/documents/guidance/UF/headache_guide.pdf [accessed 2016-05-18] [WebCite Cache ID 6hb3AC3bN]
7. Headache Classification Subcommittee of the International Headache Society. The International Classification of Headache Disorders: 2nd edition. Cephalalgia 2004;24 Suppl 1:9-160. [Medline: 14979299]
8. Regjeringen.no.: The Norwegian Ministry of Health and Care Services; 2001 Feb 08. I-12/2001 Telemedisin og ansvarsforhold URL: https://www.regjeringen.no/no/dokumenter/i-122001/id108946/ [accessed 2016-02-12] [WebCite Cache ID 6fEyFwMuS]
9. Tassorelli C, Farm I, Kettinen H, de la Torre ER, Stretenovic S, Thomas W, et al. Access to care--an unmet need in headache management? J Headache Pain 2014;15:20 [FREE Full text] [doi: 10.1186/1129-2377-15-20] [Medline: 24742114]
10. Wechsler LR. Advantages and limitations of teleneurology. JAMA Neurol 2015 Mar;72(3):349-354. [doi: 10.1001/jamaneurol.2014.3844] [Medline: 25580942]
11. Di CA, Morales-Medina JC, Palmieri B, Iannitti T. Narrative review of telemedicine consultation in medical practice. Patient Prefer Adherence 2015;9:65-75 [FREE Full text] [doi: 10.2147/PPA.S61617] [Medline: 25609928]
12. Pereira-Monteiro J, Wysocka-Bakowska M, Katsarava Z, Antonaci F, European HF. Guidelines for telematic second opinion consultation on headaches in Europe: on behalf of the European Headache Federation (EHF). J Headache Pain 2010 Aug;11(4):345-348 [FREE Full text] [Medline: 20373124]
13. Chaves-Carballo E. Diagnosis of childhood migraine by compressed interactive video. Kans Med 1992 Dec;93(12):353. [Medline: 1287284]
14. Craig J, Chua R, Wootton R, Patterson V. A pilot study of telemedicine for new neurological outpatient referrals. J Telemed Telecare 2000;6(4):225-228. [Medline: 11027124]
15. Chua R, Craig J, Wootton R, Patterson V. Randomised controlled trial of telemedicine for new neurological outpatient referrals. J Neurol Neurosurg Psychiatry 2001 Jul;71(1):63-66 [FREE Full text] [Medline: 11413265]
16. Flodgren G, Rachas A, Farmer AJ, Inzitari M, Shepperd S. Interactive telemedicine: effects on professional practice and health care outcomes. Cochrane Database Syst Rev 2015 Sep 07;9:CD002098. [doi: 10.1002/14651858.CD002098.pub2] [Medline: 26343551]
17. Elbert NJ, van Os-Medendorp H, van Renselaar W, Ekeland AG, Hakkaart-van RL, 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: 24739471]
18. Bergmo TS. How to Measure Costs and Benefits of eHealth Interventions: An Overview of Methods and Frameworks. J Med Internet Res 2015 Nov;17(11):e254 [FREE Full text] [doi: 10.2196/jmir.4521] [Medline: 26552360]
19. Statistics Norway.: Statistics Norway; 2015 Feb 19. Statistics Norway Population, 1 January 2015 URL: http://www.ssb.no/en/befolkning/statistikker/folkemengde/aar/2015-02-19?fane=tabell&sort=nummer&tabell=218738 [accessed 2015-10-02] [WebCite Cache ID 6byMJzFDM]
20. Lundqvist C, Benth JS, Grande RB, Aaseth K, Russell MB. A vertical VAS is a valid instrument for monitoring headache pain intensity. Cephalalgia 2009 Oct;29(10):1034-1041. [doi: 10.1111/j.1468-2982.2008.01833.x] [Medline: 19735531]

21. Kosinski M, Bayliss MS, Bjorner JB, Ware JE, Garber WH, Batenhorst A, Dahlöf C G H, et al. A six-item short-form survey for measuring headache impact: the HIT-6. Qual Life Res 2003 Dec;12(8):963-974. [Medline: 14651415]
22. Pasientreiser. URL: http://www.pasientreiser.no/reise-uten-rekvisisjon/dekning-av-reiseutgifter/ [accessed 2015-10-02] [WebCite Cache ID 6byN9yhsq]
23. Statistics Norway home page. 2015. Statistics Norway URL: https://www.ssb.no/en/forside;jsessionid=012850A6481FD4D54F1A7F69A637734B.kpld-as-prod03?hide-from-left-menu=true&language-code=en&menu-root-alternative-language=true [accessed 2015-10-02] [WebCite Cache ID 6byNJySUQ]
24. World Medical Association.: World Medical Association; 2008. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects URL: http://www.wma.net/en/30publications/10policies/b3/index.html [accessed 2016-05-18] [WebCite Cache ID 6hb61FDlz]
25. Poder TG, Bellemare CA, Bédard SK, Lemieux R. Social acceptance and population confidence in telehealth in Quebec. BMC Health Serv Res 2015;15:72 [FREE Full text] [doi: 10.1186/s12913-015-0727-1] [Medline: 25889230]
26. Cottrell C, Drew J, Gibson J, Holroyd K, O'Donnell F. Feasibility assessment of telephone-administered behavioral treatment for adolescent migraine. Headache 2007 Oct;47(9):1293-1302 [FREE Full text] [doi: 10.1111/j.1526-4610.2007.00804.x] [Medline: 17927645]
27. Davis LE, Coleman J, Harnar J, King MK. Teleneurology: successful delivery of chronic neurologic care to 354 patients living remotely in a rural state. Telemed J E Health 2014 May;20(5):473-477. [doi: 10.1089/tmj.2013.0217] [Medline: 24617919]
28. Trautmann E, Kröner-Herwig B. A randomized controlled trial of Internet-based self-help training for recurrent headache in childhood and adolescence. Behav Res Ther 2010 Jan;48(1):28-37. [doi: 10.1016/j.brat.2009.09.004] [Medline: 19782343]
29. Wang Y, Hunt K, Nazareth I, Freemantle N, Petersen I. Do men consult less than women? An analysis of routinely collected UK general practice data. BMJ Open 2013;3(8):e003320 [FREE Full text] [doi: 10.1136/bmjopen-2013-003320] [Medline: 23959757]
30. Lipton RB, Scher AI, Steiner TJ, Bigal ME, Kolodner K, Liberman JN, et al. Patterns of health care utilization for migraine in England and in the United States. Neurology 2003 Feb 11;60(3):441-448. [Medline: 12578925]
31. Klapper JA, Klapper A, Voss T. The misdiagnosis of cluster headache: a nonclinic, population-based, Internet survey. Headache 2000 Oct;40(9):730-735. [Medline: 11091291]
32. Rossi P, Faroni J, Tassorelli C, Nappi G. Diagnostic delay and suboptimal management in a referral population with hemicrania continua. Headache 2009 Feb;49(2):227-234. [doi: 10.1111/j.1526-4610.2008.01260.x] [Medline: 19222596]
33. de la Torre-Díez I, López-Coronado M, Vaca C, Aguado JS, de Castro C. Cost-utility and cost-effectiveness studies of telemedicine, electronic, and mobile health systems in the literature: a systematic review. Telemed J E Health 2015 Feb;21(2):81-85. [doi: 10.1089/tmj.2014.0053] [Medline: 25474190]
34. Whitten PS, Mair FS, Haycox A, May CR, Williams TL, Hellmich S. Systematic review of cost effectiveness studies of telemedicine interventions. BMJ 2002 Jun 15;324(7351):1434-1437 [FREE Full text] [Medline: 12065269]
35. Jacklin PB, Roberts JA, Wallace P, Haines A, Harrison R, Barber JA, Virtual Outreach Project Group. Virtual outreach: economic evaluation of joint teleconsultations for patients referred by their general practitioner for a specialist opinion. BMJ 2003 Jul 12;327(7406):84 [FREE Full text] [doi: 10.1136/bmj.327.7406.84] [Medline: 12855528]
36. Chua R, Craig J, Wootton R, Patterson V. Cost implications of outpatient teleneurology. J Telemed Telecare 2001;7 Suppl 1:62-64. [Medline: 11576496]

Edited by G Eysenbach; submitted 10.10.15; peer-reviewed by T Poder, F Pourmalek, I Tagliente, H Nandigam; comments to author 19.01.16; revised version received 23.02.16; accepted 05.05.16; published 30.05.16.

Please cite as:
Müller KI, Alstadhaug KB, Bekkelund SI
Acceptability, Feasibility, and Cost of Telemedicine for Nonacute Headaches: A Randomized Study Comparing Video and Traditional Consultations
J Med Internet Res 2016;18(5):e140
URL: http://www.jmir.org/2016/5/e140/
doi:10.2196/jmir.5221
PMID:27241876

Original Paper

A Remote Medication Monitoring System for Chronic Heart Failure Patients to Reduce Readmissions: A Two-Arm Randomized Pilot Study

Timothy M Hale[1,2,3*], PhD; Kamal Jethwani[1,2,3*], MPH, MD; Manjinder Singh Kandola[3], AB; Fidencio Saldana[3,4], MPH, MD; Joseph C Kvedar[1,2,3], MD

[1]Partners Healthcare, Connected Health, Boston, MA, United States

[2]Massachusetts General Hospital, Boston, MA, United States

[3]Harvard Medical School, Boston, MA, United States

[4]Brigham and Women's Hospital, Internal Medicine, Boston, MA, United States

[*]these authors contributed equally

Corresponding Author:
Timothy M Hale, PhD
Partners Healthcare
Connected Health
25 New Chardon Street Suite 300
Boston, MA,
United States
Phone: 1 617 643 9852
Fax: 1 617 225 7482
Email: tmhale@mgh.harvard.edu

Related Article:

This is a corrected version. See correction statement: https://www.jmir.org/2019/2/e13125/

Abstract

Background: Heart failure (HF) is a chronic condition affecting nearly 5.7 million Americans and is a leading cause of morbidity and mortality. With an aging population, the cost associated with managing HF is expected to more than double from US $31 billion in 2012 to US $70 billion by 2030. Readmission rates for HF patients are high—25% are readmitted at 30 days and nearly 50% at 6 months. Low medication adherence contributes to poor HF management and higher readmission rates. Remote telehealth monitoring programs aimed at improved medication management and adherence may improve HF management and reduce readmissions.

Objective: The primary goal of this randomized controlled pilot study is to compare the MedSentry remote medication monitoring system versus usual care in older HF adult patients who recently completed a HF telemonitoring program. We hypothesized that remote medication monitoring would be associated with fewer unplanned hospitalizations and emergency department (ED) visits, increased medication adherence, and improved health-related quality of life (HRQoL) compared to usual care.

Methods: Participants were randomized to usual care or use of the remote medication monitoring system for 90 days. Twenty-nine participants were enrolled and the final analytic sample consisted of 25 participants. Participants completed questionnaires at enrollment and closeout to gather data on medication adherence, health status, and HRQoL. Electronic medical records were reviewed for data on baseline classification of heart function and the number of unplanned hospitalizations and ED visits during the study period.

Results: Use of the medication monitoring system was associated with an 80% reduction in the risk of all-cause hospitalization and a significant decrease in the number of all-cause hospitalization length of stay in the intervention arm compared to usual care. Objective device data indicated high adherence rates (95%-99%) among intervention group participants despite finding no significant difference in self-reported adherence between study arms. The intervention group had poorer heart function and HRQoL at baseline, and HRQoL declined significantly in the intervention group compared to controls.

Conclusions: The MedSentry medication monitoring system is a promising technology that merits continued development and evaluation. The MedSentry medication monitoring system may be useful both as a standalone system for patients with complex medication regimens or used to complement existing HF telemonitoring interventions. We found significant reductions in risk of all-cause hospitalization and the number of all-cause length of stay in the intervention group compared to controls. Although HRQoL deteriorated significantly in the intervention group, this may have been due to the poorer HF-functioning at baseline in the intervention group compared to controls. Telehealth medication adherence technologies, such as the MedSentry medication monitoring system, are a promising method to improve patient self-management,the quality of patient care, and reduce health care utilization and expenditure for patients with HF and other chronic diseases that require complex medication regimens.

Trial Registration: ClinicalTrials.gov NCT01814696; https://clinicaltrials.gov/ct2/show/study/NCT01814696 (Archived by WebCite® at http://www.webcitation.org/6giqAVhno)

(J Med Internet Res 2016;18(4):e91) doi:10.2196/jmir.5256

KEYWORDS

heart failure; telemonitoring; telehealth; self-management; self-care; complex medication regimens; medication management; medication adherence; hospitalization length of stay; ED visits

Introduction

Heart failure (HF) affects nearly 5.7 million Americans today, contributing to 1 in 9 deaths in 2011 [1]. Furthermore, with an aging population and advancements in survival following acute cardiac injury, such as myocardial infarction, the burden of HF is expected to increase by 46% in prevalence between 2012 to 2030, resulting in approximately 8.5 million adults with HF [1,2].

In addition to its effect on morbidity and mortality, the management of HF is incredibly expensive, with aggregate costs projected to grow from US $31 billion to US $70 billion from 2012 to 2030 [2]. Heart failure remains the most frequent cause of hospitalization in patients older than 65 years [3]. Moreover, discharged patients often have subsequent complications, with 15-day readmission rates estimated at approximately 13% and 30-day readmission rates at approximately 25%; nearly half of all patients are readmitted at 6 months [4]. Unfortunately, despite efforts to improve the quality of HF care, readmission rates have not improved [4].

Heart Failure Telehealth Interventions

The burden of HF has created a sense of urgency to develop, optimize, and evaluate programs that can enable more effective transitions to self-management in patients with HF [5]. A range of new technologies (eg, high-speed Internet, wireless connections, small inexpensive sensors) enable the creation and implementation of new health care intervention strategies to improve patient self-management and achieve both more efficient care and better health outcomes [6]. Telehealth is “the application of technologies to help patients manage their own illnesses through improved self-care and access to education and support systems” [6]. Telemedicine, a closely related term, is the use of technology to deliver care at a distance to improve access, quality, and cost of care [6]. Telehealth and telemedicine programs can allow physicians and nurses to reduce face-to-face time with patients while maintaining the quality and quantity of care [7,8].

To date, most HF telehealth interventions use noninvasive remote monitoring (ie, telemonitoring) of patients’ body weight and other measures, including blood pressure, heart rate, blood oxygen saturation, and patient-reported symptoms [9]. In the advent of signs of deteriorating health, nurses communicate with physicians and patients to coordinate changes in treatment and/or schedule a visit to the clinic [9]. Research has shown that patients with chronic HF who range in age between 55 and 85 years feel confident and comfortable using telemonitoring devices [10], which allow them to experience the benefits of independence, the ability to measure and monitor their vitals, and to better manage their disease [11].

Meta-analyses of HF telemonitoring trials have been encouraging [12]. Clark et al [13] found telemonitoring programs reduced HF-related hospitalization rates by 21% and all-cause mortality by 20%. They also found significant improvement in health-related quality of life (HRQoL) in three of six studies. A meta-analysis of 13 studies found a reduction in HF-related hospitalization and all-cause mortality over 3 to 15 months follow-up [14]. Klersy et al [15] conducted a meta-analysis of 20 randomized controlled trials (RCTs) and 12 cohort studies with a median follow-up duration of 6 months in RCTs and 12 months in cohort studies. They found that telemonitoring programs were associated with lower mortality and fewer HF-related and all-cause hospitalizations. Although a large follow-up RCT conducted by Chaudhry et al [16] failed to replicate these findings, the most recent meta-analysis of HF telemonitoring studies conducted in 2015 found improved outcomes [17] and a number of recent HF telemonitoring trials suggest the possibility and promise of these technologies to improve care and clinical outcomes [18-26].

Although many HF telemonitoring interventions have been found to be useful immediately following hospitalization, their long-term impacts remain to be fully characterized [9] and the benefits may decrease over time. For example, Clarke et al [14] found that the reduced risk of hospitalization seen at 12 months postdischarge was no longer significant at 15 months. Agboola et al [18] observed a similar pattern with decreased hospitalizations at 30 days and at the end of the 4-month program. However, there was no observed reduction in hospitalizations at 12 months. Therefore, there is a need to find alternative telemonitoring strategies and devices that are effective in helping HF patients to self-manage their medical condition on a long-term basis.

Medication Adherence

One key component of HF patient self-management is medication adherence. Poor medication adherence contributes to poor quality of life [27] and increased risk of mortality and morbidity [28], hospitalization [28,29], and high health care costs [2] among HF patients. Adherence ranging from 2% to 90% has been reported [30,31] with a typical adherence of approximately 40% to 60% [28,32]. It is estimated that poor adherence among HF patients contributes to approximately US $300 billion in health care costs each year [33].

There is widespread agreement that efforts to improve medication adherence and patient self-care are needed to improve HF patients' quality of life and reduce the risk of hospitalization [34,35]. However, achieving high medication adherence is difficult in this population for several reasons. First, adequate adherence for HF patients may be higher than what is needed to manage other medical conditions. Adequate adherence is often defined as 80% or greater, but Wu et al [36] found that daily dose adherence of 88% or greater was required to achieve longer time to first emergency department (ED) visit and decreases in hospitalizations and mortality among HF patients. Second, HF patients are often prescribed a complex medication regimen consisting of multiple medications to achieve symptom control, reduce morbidity and mortality, and to improve quality of life [37,38]. Finally, cognitive deficits and memory problems, which are more common among older adults with HF [39,40], make it difficult for patients to remember to take their medications [29]. Riegal and Knafl [29] found that impaired cognition was an important factor interacting with poor adherence and recommend using devices to remind patients to take their medications. Therefore, a solution is needed to help HF patients better manage their medications and remember to take them as prescribed long after traditional telemonitoring interventions end.

Medication Adherence Telehealth Interventions

An electronic remote medication monitoring system could prove efficacious in aiding HF patients to adhere to their complicated medication regimen. Evidence suggests that patient-focused adherence interventions are more effective in improving adherence rates than efforts directed at health care provider behaviors [32]. In fact, many recent remote monitoring interventions have sought to specifically improve medication adherence through diverse means, such as increased patient education and health literacy, provider and pharmacist consultation, phone-based adherence assessments and positive behavior encouragement, and electronic reminders. However, many of these are early phase trials and results are so far met with cautious optimism. Although many suggest usability and patient satisfaction, as well as improved adherence measures, their ultimate effect on health outcomes remains largely unmeasured and requires further evaluation [32,41,42].

Study Goals

The primary goal of this randomized controlled pilot study is to compare the MedSentry medication monitoring system versus usual care in older HF adults in patients who recently completed a HF telemonitoring program. The MedSentry medication monitoring system is a novel technology consisting of two parts: (1) a remotely monitored electronic pillbox that alerts people when it is time to take their medications and (2) a monitoring center with advisors who contact patients and caregivers when medications are not taken. We hypothesized that remote medication monitoring would be associated with fewer unplanned hospitalizations and ED visits, increased medication adherence, and improved HRQoL versus usual care. We also examined user satisfaction and usability with the MedSentry medication monitoring system among participants in the intervention arm.

Methods

Study Design

This pilot study was conducted as a RCT. Participants randomized to the intervention group used the MedSentry medication monitoring system for medication management. Participants randomized to the control group continued to use their usual medication reminder method. The intervention period was 90 days. Participants were given the option to enroll in a second 90-day period in which participants in the control group used the MedSentry medication monitoring system and participants in the intervention arm returned to their usual medication reminder method. The results reported in this paper examine only the first 90-day intervention period. The initial goal was to enroll a total of 70 participants with 35 randomized to each of the two study arms. However, due to slow enrollment, the study was ended early with 29 participants who completed enrollment and randomization.

Study Participants

Participants were recruited from a list of HF patients who had successfully completed a HF telemonitoring program at Massachusetts General Hospital (MGH) or Brigham and Women's Hospital (BWH) and had been hospitalized in the previous 24 months. Patient lists were reviewed by study staff to identify potential participants based on other inclusion/exclusion criteria and, if met, were approved by the MGH- or BWH-affiliated physician for participation in the study.

The inclusion criteria for this study were (1) take at least three and no more than 10 different daily medications, (2) take medications no more than four specified times each day (ie, morning, afternoon, early evening, bedtime), (3) able to sort and manage their own medications, (4) have a telephone or cell phone, (5) live in the greater Boston area, and (6) speak, read, and write English. Patients were excluded if they met any of the following: (1) vision or hearing impaired (ie, unable to hear an alarm similar to a clock alarm or oven alarm), (2) dementia or other conditions precluding the participant from providing informed consent, (3) awaiting revascularization, cardiac resynchronization, or heart transplant, and (4) terminal illness.

Eligible patients were screened by phone to confirm eligibility and be informed of the study. Due to the difficulty in scheduling the device installation, patients who agreed to participate were randomized during the screening phone call. Participants assigned to the control group provided consent over the phone

and were mailed a copy of the consent form and enrollment questionnaire to complete. Participants assigned to the intervention group were scheduled for a study visit in which they completed the informed consent process and were trained in the use of the MedSentry device. The Partners HealthCare Institutional Review Board approved all study procedures.

The MedSentry Medication Management System

The MedSentry medication monitoring system consists of two parts: (1) a remotely monitored electronic device ("device") that alerts participants when it is time to take their medications and (2) a monitoring center with advisors who contact participants and caregivers when medications are not taken. The device is installed in the participant's home and data are transmitted to the monitoring center via the Internet.

The device is approximately the size of a small microwave oven. The top of the device consists of a series of small, removable bins arranged in a 7 by 4 configuration (seven days of the week and four medication times per day). A lid on the top of each bin detects when a bin is opened. The bottom of each bin is clear plastic. Cameras located under the bins transmit an image of the contents to the monitoring center. During the installation, participants were trained in how to refill the device with the medications they were prescribed for each time of day. The monitoring center was provided a medication list and an image of the correctly loaded medications for reference.

The MedSentry medication monitoring system uses several methods to ensure participants take their medications as prescribed. First, the device provides a visual cue (blue lights around a bin) and an audio alarm to alert a participant when it is time to take their medication. If a dose is not taken within 30 minutes, an advisor at the monitoring center calls the participant. After three attempts over a 45-minute time span to contact the participant, a voice message is left and a call is placed to an optional caregiver who has agreed to be contacted and to follow up with the participant.

Participants were responsible for refilling the device and communicating medication changes to the monitoring center. Each week, the monitoring center advisor called to remind participants to refill the device. When participants refilled the device, a new image was transmitted and compared at the monitoring center. Participants were contacted by phone to correct refill errors. When there was a change in medications, participants used the "call request button" on the front of the device to notify the monitoring center of changes. The monitoring center verified that the pill tray was filled correctly and updated the medication list and reference image. Monthly medication adherence reports were sent to the participant's physician and caregiver, if requested.

Data Collection and Outcome Measures

Participants in both study arms completed enrollment and closeout questionnaires (Multimedia Appendix 1). Participants in the control arm completed and returned questionnaires by mail. Participants in the intervention arm completed questionnaires at the installation visit and at closeout when the device was removed from their home. The enrollment questionnaire collected data on demographics and baseline technology use. The closeout questionnaire for the intervention arm participants included questions to assess satisfaction and usefulness of the MedSentry device and monitoring services and their willingness to recommend the device.

The enrollment and closeout questionnaires contained items to assess changes in medication adherence, health, and HRQoL. Medication adherence was assessed using a self-reported measure and data collected by the device. Self-reported adherence was assessed using a single question from the Medical Outcomes Study (MOS) [43,44]. Participants were asked, "How often did you take your medications as prescribed (on time without skipping doses) during the past 4 weeks?" A dichotomous measure was created to indicate "nonadherent" (none, a little, or some of the time) or "adherent" (most or all the time). A direct measure of medication adherence was derived using data from the MedSentry device and monitoring center logs. Adherence for each 30-day period was calculated using data on the number of pills taken divided by the number of pills prescribed. If medications were not taken within 1 hour of the prescribed time, they were coded as "missed" and coded as "taken" if confirmed by an outreach phone call from the monitoring center.

Health was assessed using a single item on self-reported general health status (1=poor to 5=excellent). Depression was assessed using the 8-item Patient Health Questionnaire (PHQ-8). Each of the PHQ-8 items is scored from 0 to 3 to generate a total score from 0 to 24. Generally accepted cut points for depression severity are 5-9=mild, 10-14=moderate, 15-19=moderately severe, and 20-24=severe [45]. The PHQ-8 has been found to be a valid and reliable measure of depression among HF patients [46]. To assess baseline heart function the New York Heart Association (NYHA) Functional Classification was extracted from participants' electronic medical records for the clinic visit nearest the study enrollment date. Heart functioning is classified as one of four levels, ranging from Class I="no symptoms with ordinary activity" to Class IV="unable to carry out any physical activity without discomfort; symptoms of cardiac insufficiency may be present even at rest." When no classification was recorded in the electronic medical record (EMR), the study research nurse created a classification based on the clinical notes recorded nearest the study enrollment date.

Health-related quality of life was measured using the Minnesota Living with Heart Failure Questionnaire (MLHFQ) [47]. The questionnaire consists of 21 items that assess the impact of HF and HF treatment on key physical, emotional, and social dimensions of a patient's life during the past four weeks. Responses are coded from 0=does not apply and 1=very little to 5=very much. In addition to a total summary score, a physical subscore and emotional subscore can be created [47,48]. The instrument has good construct validity and test-retest reliability [47,49].

Data on hospitalization and ED visits came from two sources. The primary source was the participants' EMR. The second source was a series of questions on the closeout questionnaire about the type and timing of hospitalization and ED visits during the 90-day study period. Our analysis focused on unplanned hospitalizations; therefore, planned hospitalizations were

excluded from our calculations of hospital visits and days hospitalized.

Data Analysis

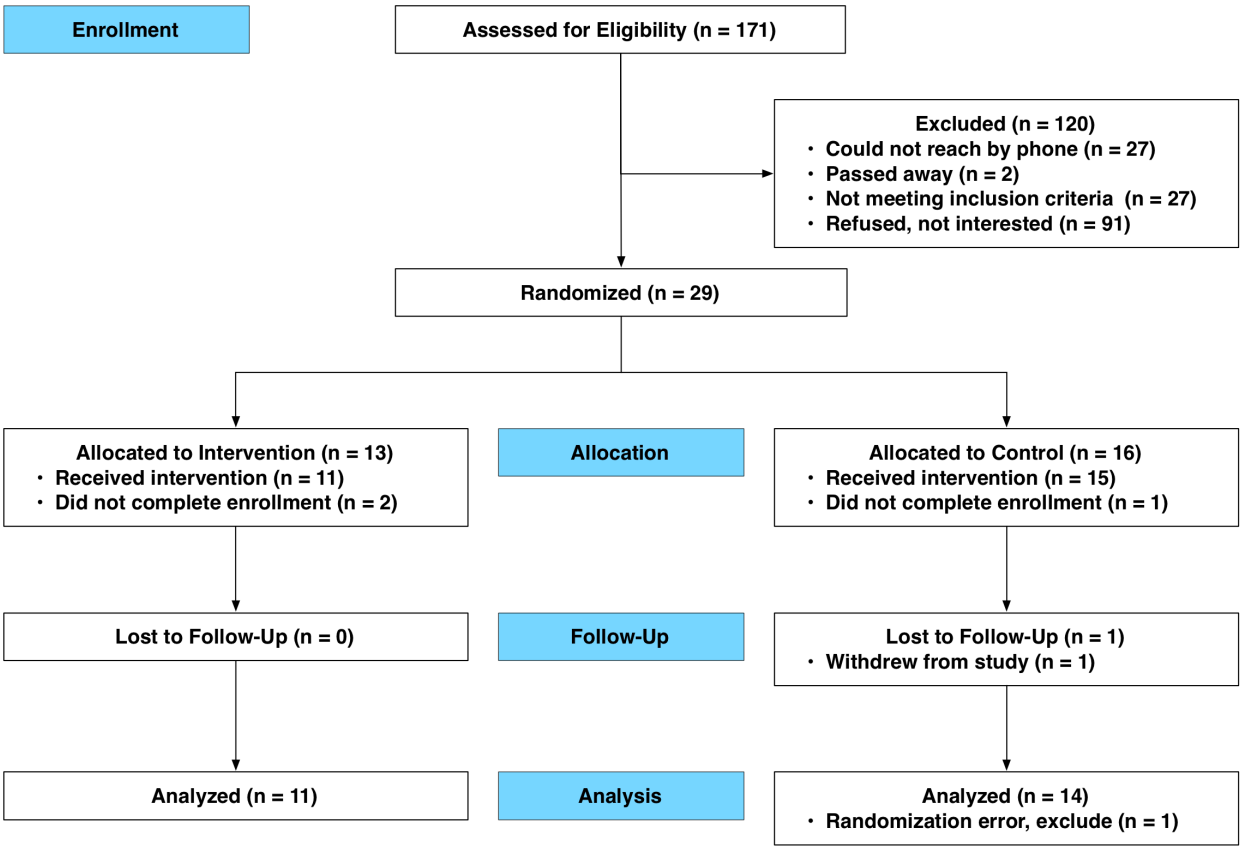
Baseline characteristics were summarized and compared for both study arms using percentages for categorical variables and means and standard deviations for continuous variables. Comparisons on outcome measures between intervention and control arms at the close of the study were conducted using unpaired *t* tests, unpaired proportion test, and the Fisher exact test. For skewed data (ie, ED visits and days hospitalized), Wilcoxon rank sum tests were conducted. Descriptive statistics were used to summarize results on usability and satisfaction with the MedSentry medication monitoring system. Data analysis was performed with Stata 14 with an alpha of .05 set a priori .

Results

Study Flow

Of the 171 patients assessed for eligibility, 29 were randomized (see Figure 1). Thirteen participants were randomized to the intervention arm, two participants did not complete the enrollment process, and 11 completed the study. In total, 16 participants were randomized to the control group; however, one did not complete the enrollment process and one was mistakenly randomized to the control group and was excluded from analysis. Of the 14 enrolled participants, one withdrew leaving 13 participants who completed the study. Conducted as an intention-to-treat analysis, we included the participant who withdrew from the study in the analysis of hospitalizations and ED visits.

Figure 1. Participant enrollment and inclusion.



Baseline Characteristics

Descriptive statistics for demographics and baseline characteristics are presented in Table 1. The mean age of participants was 71.7 years (SD 11.2) and 64% (16/25) were male. Participants were predominantly white and had one or more years of college education. Overall, 44% (11/25) were married, 40% (10/25) lived alone in the household, 64% (16/25) had used the Internet, and 72% (18/25) had a cellphone. There were no statistically significant differences in these characteristics between participants in the intervention and control arms. Of the 25 participants, 18 (72%) reported they currently used a medication reminder tool or method. Seventy-two percent (18/25) of participants were categorized as “adherent” based on self-reported adherence.

Table 1. Baseline characteristics of participants and comparison by study arm.

Variables	Total (N=25)	Control (n=14)	Intervention (n=11)	*P*
Sociodemographics				
Age (years), mean (SD)	71.7 (11.2)	74.4 (10.4)	68.4 (11.8)	.19
Gender (male), n (%)	16 (64)	9 (64)	7 (64)	.97
Race (white), n (%)	22 (88)	13 (93)	9 (82)	.40
Married, n (%)	11 (44)	7 (50)	4 (36)	.50
Lives alone, n (%)	10 (40)	6 (43)	4 (36)	.74
Education (≥1 years of college), n (%)	18 (72)	11 (79)	7 (63)	.41
Technology use and attitudes				
Internet user, n (%)	16 (64)	10 (71)	6 (55)	.38
Mobile phone ownership, n (%)				
No mobile phone	3 (7)	1 (7)	2 (18)	.69
Cellphone	18 (72)	11 (79)	7 (64)	
Smartphone	4 (16)	2 (14)	2 (18)	
Medication adherence				
MOS-Adhere (adherent), n (%)	18 (72)	11 (79)	7 (64)	.41
Using a medication reminder tool, n (%)	18 (72)	11 (79)	7 (64)	.41
Health status				
NYHA Functional Classification, n (%)				
Class I	12/24 (50)	11/13 (85)	1 (9)	.001
Class II	10/24 (42)	1/13 (8)	9 (82)	
Class III	2/24 (8)	1/13 (8)	1 (9)	
Class IV				
Self-rated health, n (%)				
Poor	1 (4)	0 (0)	1 (9)	.68
Fair	11 (44)	6 (43)	5 (46)	
Good	10 (40)	5 (36)	5 (46)	
Very good	2 (8)	2 (14)	0 (0)	
Excellent	1 (4)	1 (7)	0 (0)	
Depression (PHQ-8), mean (SD)	5.0 (4.8)	3.5 (4.6)	6.8 (4.5)	.08
Health-related quality of life (MLHFQ), mean (SD)[a]				
MLHFQ total score	34.3 (25.6)	26.2 (23.1)	43.7 (25.9)	.10
MLHFQ physical score	15.7 (12.7)	10.8 (11.4)	21.9 (11.8)	.03
MLHFQ emotional score	6.5 (6.7)	5.9 (6.2)	7.2 (7.7)	.65

[a]The MLHFQ is scored so that higher values indicate an adverse impact on quality of life.

At baseline, intervention participants appeared to have been in significantly worse health than participants assigned to the control arm. Intervention participants had more severe HF symptoms, as assessed using the NYHA Functional Classification. In the intervention arm, 9% (1/11) were Class I and 82% (9/11) were Class II compared to controls, which were 85% (11/13) Class I and 8% (1/13) Class II (*P*=.001). The MLHFQ physical subscale was significantly worse in the intervention group compared to controls (mean 21.9, SD 11.8 vs mean 10.8, SD 11.4, *P*=.03). There were no other statistically significant differences in baseline characteristics.

Emergency Department Visits and Hospitalizations

Table 2 presents the number of participants who had one or more ED visits or hospitalizations. There was no statistically significant difference between study arms in ED visits or for HF-related and non-HF-related hospitalizations. There were significantly fewer all-cause hospitalizations among the intervention group compared to controls. Approximately 9%

(1/11) of intervention participants were hospitalized one or more times compared to 50% (7/14) of control participants (P=.04), a relative risk reduction in hospitalization of approximately 82%. Additional analysis using logistic regression was conducted to control for baseline differences in NYHA classification and the MLHFQ physical subscale. However, the addition of these baseline measures did not improve the fit of the model and were not significant.

Table 2. Emergency department visits and hospitalizations (one or more occurrences during the 90-day study period) by study arm.

Variables	Control, n (%) (n=14)	Intervention, n (%) (n=11)	*P*
Emergency department visits			
Heart failure-related	3 (21)	1 (9)	.60
Non-heart failure-related	4 (29)	3 (27)	>.99
All cause[a]	6 (43)	3 (27)	.68
Hospitalizations			
Heart failure-related	4 (29)	1 (9)	.34
Non-heart failure-related	4 (29)	1 (9)	.34
All cause[a]	7 (50)	1 (9)	.04

[a]Some participants reported both HF- and non-HF-related ED visit or hospitalization; therefore, the percentage of participants with one or more all-cause ED visits or hospitalizations is lower than the sum of the two types.

Table 3 presents the results comparing the total number of ED visits, hospitalizations, and length of stay if admitted to the hospital. The number of ED and hospitalization visits did not differ significantly between study arms. The intervention arm had significantly fewer all-cause hospitalization days compared to the controls (4 vs 34, P=.03) and there was a reduction in the number of days for HF-related and non-HF-related hospitalizations, but this did not meet statistical significance (P=.24).

Table 3. Emergency department visits, hospitalizations, and hospitalization length of stay for both study arms at closeout.

Variables	Control (n=14)		Intervention (n=11)		*P*
	n	Mean (SD)	n	Mean (SD)	
Emergency department visits					
All cause	7	0.50 (0.65)	4	0.36 (0.67)	.50
Heart failure	3	0.21 (0.43)	1	0.09 (0.30)	.41
Non-heart failure	4	0.29 (0.47)	3	0.27 (0.46)	.94
Hospitalizations					
All cause	8	0.57 (0.65)	2	0.18 (0.60)	.06
Heart failure	4	0.29 (0.47)	1	0.09 (0.30)	.24
Non-heart failure	4	0.29 (0.47)	1	0.09 (0.30)	.24
Hospitalization length of stay (days)					
All cause	34	2.43 (3.13)	4	0.36 (1.21)	.03
Heart failure	19	1.36 (2.68)	2	0.18 (0.60)	.20
Non-heart failure	15	1.07 (2.06)	2	0.18 (0.60)	.21

Medication Adherence

Table 4 presents the results of tests for differences between the intervention and control arms on medication adherence, health status, and HRQoL. At study closeout, there was no significant difference in self-reported medication adherence. The MOS single-item measure indicated that 67% of participants in the two study arms were adherent. Objective data on adherence generated by the MedSentry device used in the intervention group indicated monthly adherence levels of 94% to 99% after efforts made by the monitoring center staff to confirm medications were taken as prescribed.

Table 4. Medication adherence, health status, and health-related quality of life for both study arms at closeout.[a]

Variables	Control (n=13)	Intervention (n=11)	*P*
Medication adherence			
MOS-Adhere (adherent), n (%)	9 (69)	8 (73)	.61
Adherence, MedSentry device data, mean % (SD)			
Total (90 days)		96.8 (7.2)	
Month 1 (days 1-30)		98.7 (2.1)	
Month 2 (days 31-60)		97.4 (5.0)	
Month 3 (days 61-90)		94.2 (11.2)	
Health status			
Self-rated health			
Poor	0/12 (0)	1 (9)	.82
Fair	5/12 (42)	3 (27)	
Good	6/12 (50)	6 (55)	
Very Good	1/12 (8)	1 (9)	
Excellent	0/12 (0)	0 (0)	
Depression (PHQ-8), mean (SD)	3.5 (3.6)	4.5 (2.6)	.46
Health-related quality of life (MLHFQ), mean (SD)[b]			
MLHFQ total score	28.2 (22.3)	62.2 (20.6)	.002
MLHFQ physical score	11.2 (10.8)	29.8 (10.7)	.001
MLHFQ emotional score	5.6 (6.0)	11.5 (6.7)	.03

[a]Missing cases for some comparisons is because of incomplete responses on the closeout questionnaire.

[b]The MLHFQ is scored so that higher values indicate an adverse impact on quality of life.

Health Status and Health-Related Quality of Life

There was no significant difference in self-rated health or in depression assessed using the PHQ-8. However, the intervention arm had significantly worse HRQoL as measured using the MLHFQ and on the two subscales that assessed physical and emotional dimensions. For example, intervention arm participants had a mean MLHFQ score of 62.2 (SD 20.6) and the control arm had a mean of 28.2 (SD 22.3, *P*=.002).

Usefulness and Satisfaction

A series of questions on the closeout questionnaire were used to assess the intervention participants' ratings of the usefulness of 10 MedSentry medication monitoring system features (see Table 5). At least 50% of participants ranked all 10 features as "mostly" or "extremely" useful. The features rated most useful (based on number of "extremely" useful ratings) were the scan that ensured the device was loaded correctly, receiving a call if the wrong medicine was taken, and lights indicating which bin to open. The two features participants rated as least useful were the medication adherence reports sent to the participants' physician and caregiver and the arrows on the device control panel. Two questions were used to assess whether participants would recommend or want to continue to use the MedSentry medication monitoring system if it were available (results not shown); 70% (7/10) strongly agreed that they would recommend the MedSentry medication monitoring system to a friend or family member and 50% (5/10) strongly agreed they would like to continue to use the MedSentry medication monitoring system if it was made available.

Table 5. Usefulness of MedSentry medication monitoring system features (n=11).[a]

Usefulness items	Did not use	Not at all	A little	Mostly	Extremely
Scan to ensure meds loaded correctly	0 (0)	1 (9)	1 (9)	1 (9)	8 (73)
Receiving a call if removed wrong meds	1 (9)	0 (0)	2 (18)	1 (9)	7 (64)
Glowing bins for meds to take	0 (0)	0 (0)	3 (27)	1 (9)	7 (64)
Alarm if wrong bin opened	1 (9)	0 (0)	1 (9)	3 (27)	6 (55)
Receiving a call if missed meds	0 (0)	1 (9)	2 (18)	2 (18)	6 (55)
Ring tone reminder to take meds	0 (0)	0 (0)	2 (18)	3 (27)	6 (55)
Reminder messages on control panel	1 (9)	2 (18)	1 (9)	2 (18)	5 (46)
Call button to request a call	1 (9)	1 (9)	2 (18)	3 (27)	4 (36)
Reports sent to doctor and caregiver	4 (36)	0 (0)	0 (0)	3 (27)	4 (36)
The arrows on the control panel (n=6)	3 (50)	0 (0)	0 (0)	1 (17)	2 (33)

[a]Missing cases for some comparisons was because of incomplete responses on the closeout questionnaire. Percentages may not total to 100 due to rounding error.

Discussion

This study explored the effects of the MedSentry medication monitoring system, a remote medication management system among HF patients. We hypothesized that remote medication monitoring would be associated with fewer hospitalizations and ED visits, increased medication adherence, and improved HRQoL compared to usual care. We further sought to assess user satisfaction and usability among participants in the intervention arm.

The first hypothesis was that the MedSentry medication monitoring system would be associated with fewer hospitalizations and ED visits compared to usual care. There were a significantly lower number of all-cause hospitalizations in the intervention arm relative to the controls. In the intervention arm, approximately 9% (1/11) were hospitalized one or more times compared to 50% (7/14) in the control arm, a reduction in the relative risk of approximately 82%. Furthermore, there were a significantly lower number of all-cause days hospitalized in the intervention arm compared to controls and fewer HF and non-HF days hospitalized in the intervention arm compared to controls, although this difference did not reach statistical significance. However, there was no statistically significant difference in the number of ED visits.

The reduction in the days hospitalized is encouraging given the small sample size of this pilot study and the relatively poorer health of the intervention participants compared to the controls. Previous studies examining the effects of HF telemonitoring interventions have been mixed. A recent large scale RCT found no significant reduction in hospitalization and ED visits [16], whereas systematic reviews [15,17] and smaller studies have found positive effects [19,22-26]. This variation is likely due to differences in the type of telemonitoring intervention, inclusion criteria and variations in the health of patients enrolled in the programs, and real-world differences in implementation of research studies across multiple clinics and sites [9]. A recent analysis of a standard-of-care HF telemonitoring program, combined with telephone nursing support, was found to be associated with a reduction in hospitalization rates and mortality up to 120 days postdischarge compared to controls, which marked the end of the telemonitoring program [18]. However, 8 months later there was no difference between those who participated in the program and matched controls. Thus, there is reason to believe that telemonitoring systems can substantially reduce health care utilization following HF-related hospitalization if systems are well integrated with existing care delivery and targeted toward patients at risk. Additionally, existing telemonitoring systems could be complemented by other interventions strategies to extend telemonitoring beyond 90 to 120 days, which is typical of most programs.

Objectively measured medication adherence in the intervention group ranged from 94% to 99%, as measured by data generated directly by the devices. However, the intervention did not improve medication adherence as measured by self-report (ie, MOS). The observed lack of improvement may be due to the lack of reliability in using self-reports as a measure of true adherence. For example, in a previous study, researchers found similar self-reported adherence rates (ie, 72%) but objectively measured adherence as low as 54% when comparing the single-item MOS adherence measure to objective measurements using a medication event monitoring system (MEMS) among HF patients [28]. Thus, it is possible that participants overestimated their level of adherence at baseline and may have underreported adherence rates at closeout because they were alerted each time they erred in taking or refilling their medication. In fact, over the course of the study, the monitoring center detected weekly refill errors of 62%, which required calls to participants to correct. Second, at baseline, 72% of patients were using some form of medication reminder system and all participants had completed a telemonitoring intervention.

In our review of the literature, we were unable to find studies evaluating telemonitoring interventions to improve medication adherence for HF patients or patients with complex medication regimens. However, a meta-analysis of other types of medication adherence interventions for HF patients' found, overall, that programs have a modest effect, especially when the focus is on patients and their medication-taking behaviors [32]. This meta-analysis also found that a focus on interventions aimed at

modifying one patient behavior are more effective in increasing adherence than interventions addressing multiple behaviors. This suggests that innovative telemedicine medication adherence interventions may prove especially effective due to the small number of new behaviors (ie, refilling the device) required by patients and the support provided by the monitoring center staff to guide patients in correcting refill errors, which were found to be considerable in number.

The third hypothesis was that the MedSentry medication monitoring system would be associated with improved HRQoL. In contrast, we found the opposite effect—participants in the intervention arm had a significantly poorer HRQoL at closeout than the control arm as measured using the MLHFQ physical subscale. However, this deterioration in HRQoL might be due to differences in HF status between the two study arms at enrollment. Participants in the intervention arm had significantly poorer HF-related health, as classified by the NYHA Functional Classification and the MLHFQ physical subscale. Thus, intervention arm participants appear to have more severe HF than the control arm participants and this may have contributed to a greater decline in health and HRQoL compared to control arm participants over the course of the study.

Finally, this study assessed usability and patient satisfaction related to using the MedSentry medication monitoring system. Users rated the device highly, with at least half of all participants ranking each listed feature as “mostly” or “extremely” useful, and a majority strongly agreeing they would recommend the MedSentry medication monitoring system to a friend or family member. Patients liked the features of the device that promoted supervised engagement, such as the light indicating which bin to open and the scan to ensure correct loading. Although users liked the idea of receiving a call if they took the wrong medication, they were less amenable to the idea of adherence reports being sent to their doctor and caregiver.

Limitations

Despite the strengths of using a randomized controlled study design to evaluate the effect of the intervention, there are limitations to this study that should be considered when interpreting results. First, the size of the sample is small and the power to detect differences at baseline and closeout is low. Unfortunately, recruitment was slow and the study was ended early before achieving the original goal of 35 participants per study arm. Future research is needed with a larger sample size to ensure statistical power to evaluate primary and secondary outcomes. One reason many patients declined to participate, or were not approved by physicians, was due to a belief they were managing their medications adequately. Research shows this is often not the case and efforts to further educate providers and patients about this fact may improve enrollment rates.

A second limitation is that self-reported measures of medication adherence are not the best method to assess the effect of remote medication management systems on adherence. When possible, a study design using a MEMS to objectively track medication adherence in the control arm would be preferable to self-reports. Unfortunately, this is difficult to accomplish because no devices exist that make it easy to track the multiple medications that most HF patients have been prescribed. One alternative might be to use a “dummy” device that is similar in design to the test device, but lacks the reminders and alerts and is used only for tracking medication adherence.

Finally, a third limitation is that participants in our two study arms differed significantly in NYHA classification and HRQoL. We found that controlling for these differences using logistic regression models did not change our results. However, future research using a larger sample and/or inclusion/exclusion criteria that is more narrowly defined regarding HF condition at enrollment would be useful in creating a sample that is better balanced on baseline characteristics.

Conclusion

The MedSentry medication monitoring system is a promising technology that merits continued development and evaluation. Most existing telemedicine HF interventions monitor vital signs and self-reported symptoms for 30 to 90 days postdischarge, whereas the MedSentry medication monitoring system provides a relatively low-cost means to remotely monitor HF patients’ medication adherence. The home device is used to remind patients when they should take their medications and the innovative use of cameras to monitor the contents of each medication bin enables advisors at the remote monitoring center to follow-up with patients by phone if there are missed medications or refill errors. As a standalone system, we found that the use of the MedSenty medication monitoring system was associated with an 80% reduction in the risk of unplanned all-cause hospitalization and a significant decrease in the number of unplanned all-cause days hospitalized in the intervention arm compared to usual care. The MedSentry medication monitoring system may be useful both as a standalone system for patients with complex medication regimens or used to complement existing HF telemonitoring interventions. Telemonitoring medication adherence technologies, such as the MedSentry medication monitoring system, are a promising method to improve the quality of patient care and reduce health care utilization and expenditure for patients with HF and other chronic diseases that require complex medication regimens.

Acknowledgments

The authors would like to acknowledge Ellen Murachver, MSW, and Caitlin Masters, MPH, for their contributions to carrying out the study; Victoria Powell, MPH, for her help in data coding; and to Stephen Agboola, MD, MPH, for suggestions on the manuscript and data analyses. We also want to thank Stephanie A Moore, MD, for her help in identifying and recruiting participants for this study. Funding for this investigator-initiated study was provided by Presentcare, Inc.

Conflicts of Interest

None declared

Multimedia Appendix 1

Enrollment and Closeout Questionnaires.

[PDF File (Adobe PDF File), 1MB - jmir_v18i5e91_app1.pdf]

References

1. Mozaffarian D, Benjamin E, Go A. Heart disease and stroke statistics--2015 update: a report from the American Heart Association. Circulation 2015 Jan 27;131(4):e29-e322.
2. Heidenreich PA, Albert NM, Allen LA, Bluemke DA, Butler J, Fonarow GC, American Heart Association Advocacy Coordinating Committee, Council on Arteriosclerosis, ThrombosisVascular Biology, Council on Cardiovascular RadiologyIntervention, Council on Clinical Cardiology, Council on EpidemiologyPrevention, Stroke Council. Forecasting the impact of heart failure in the United States: a policy statement from the American Heart Association. Circ Heart Fail 2013 May;6(3):606-619 [FREE Full text] [doi: 10.1161/HHF.0b013e318291329a] [Medline: 23616602]
3. Lloyd-Jones D, Adams R, Carnethon M, De Simone G, Ferguson TB, Flegal K, American Heart Association Statistics Committee and Stroke Statistics Subcommittee. Heart disease and stroke statistics--2009 update: a report from the American Heart Association Statistics Committee and Stroke Statistics Subcommittee. Circulation 2009 Jan 27;119(3):e21-e181 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.108.191261] [Medline: 19075105]
4. Ross JS, Chen J, Lin Z, Bueno H, Curtis JP, Keenan PS, et al. Recent national trends in readmission rates after heart failure hospitalization. Circ Heart Fail 2010 Jan;3(1):97-103 [FREE Full text] [doi: 10.1161/CIRCHEARTFAILURE.109.885210] [Medline: 19903931]
5. Yancy CW, Jessup M, Bozkurt B, Butler J, Casey DE, Drazner MH, American College of Cardiology Foundation, American Heart Association Task Force on Practice Guidelines. 2013 ACCF/AHA guideline for the management of heart failure: a report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines. J Am Coll Cardiol 2013 Oct 15;62(16):e147-e239 [FREE Full text] [doi: 10.1016/j.jacc.2013.05.019] [Medline: 23747642]
6. Kvedar J, Coye MJ, Everett W. Connected health: a review of technologies and strategies to improve patient care with telemedicine and telehealth. Health Aff (Millwood) 2014 Feb;33(2):194-199. [doi: 10.1377/hlthaff.2013.0992] [Medline: 24493760]
7. Kedar I, Ternullo JL, Weinrib CE, Kelleher KM, Brandling-Bennett H, Kvedar JC. Internet based consultations to transfer knowledge for patients requiring specialised care: retrospective case review. BMJ 2003 Mar 29;326(7391):696-699 [FREE Full text] [doi: 10.1136/bmj.326.7391.696] [Medline: 12663408]
8. Vasquez MS. Down to the fundamentals of telehealth and home healthcare nursing. Home Healthc Nurse 2008 May;26(5):280-7; quiz 288. [doi: 10.1097/01.NHH.0000318943.62538.0a] [Medline: 18469601]
9. Anker SD, Koehler F, Abraham WT. Telemedicine and remote management of patients with heart failure. Lancet 2011 Aug 20;378(9792):731-739. [doi: 10.1016/S0140-6736(11)61229-4] [Medline: 21856487]
10. Nahm E, Blum K, Scharf B, Friedmann E, Thomas S, Jones D, et al. Exploration of patients' readiness for an eHealth management program for chronic heart failure: a preliminary study. J Cardiovasc Nurs 2008;23(6):463-471. [doi: 10.1097/01.JCN.0000338930.89210.79] [Medline: 18953206]
11. Lehmann C, Mintz N, Giacini J. Impact of telehealth on healthcare utilization by congestive heart failure patients. Dis Manag Health Out 2006;14(3):163-169. [doi: 10.2165/00115677-200614030-00005]
12. Inglis SC, Clark RA, McAlister FA, Ball J, Lewinter C, Cullington D, et al. Structured telephone support or telemonitoring programmes for patients with chronic heart failure. Cochrane Database Syst Rev 2010(8):CD007228. [doi: 10.1002/14651858.CD007228.pub2] [Medline: 20687083]
13. Clark RA, Inglis SC, McAlister FA, Cleland JG, Stewart S. Telemonitoring or structured telephone support programmes for patients with chronic heart failure: systematic review and meta-analysis. BMJ 2007 May 5;334(7600):942 [FREE Full text] [doi: 10.1136/bmj.39156.536968.55] [Medline: 17426062]
14. Clarke M, Shah A, Sharma U. Systematic review of studies on telemonitoring of patients with congestive heart failure: a meta-analysis. J Telemed Telecare 2011 Jan;17(1):7-14. [doi: 10.1258/jtt.2010.100113] [Medline: 21097564]
15. Klersy C, De Silvestri A, Gabutti G, Regoli F, Auricchio A. A meta-analysis of remote monitoring of heart failure patients. J Am Coll Cardiol 2009 Oct 27;54(18):1683-1694 [FREE Full text] [doi: 10.1016/j.jacc.2009.08.017] [Medline: 19850208]
16. Chaudhry SI, Mattera JA, Curtis JP, Spertus JA, Herrin J, Lin Z, et al. Telemonitoring in patients with heart failure. N Engl J Med 2010 Dec 9;363(24):2301-2309 [FREE Full text] [doi: 10.1056/NEJMoa1010029] [Medline: 21080835]
17. 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 Mar;17(3):e63 [FREE Full text] [doi: 10.2196/jmir.4174] [Medline: 25768664]

18. Agboola S, Jethwani K, Khateeb K, Moore S, Kvedar J. Heart failure remote monitoring: evidence from the retrospective evaluation of a real-world remote monitoring program. J Med Internet Res 2015;17(4):e101 [FREE Full text] [doi: 10.2196/jmir.4417] [Medline: 25903278]
19. Andrikopoulou E, Abbate K, Whellan DJ. Conceptual model for heart failure disease management. Can J Cardiol 2014 Mar;30(3):304-311. [doi: 10.1016/j.cjca.2013.12.020] [Medline: 24565255]
20. Gorthi J, Hunter C, Mooss A, Alla V, Hilleman D. Reducing heart failure hospital readmissions: a systematic review of disease management programs. Cardiol Res 2014 Jan;5(5):126-138. [doi: 10.14740/cr362w]
21. Koehler F, Winkler S, Schieber M, Sechtem U, Stangl K, Böhm M, et al. Telemedical Interventional Monitoring in Heart Failure (TIM-HF), a randomized, controlled intervention trial investigating the impact of telemedicine on mortality in ambulatory patients with heart failure: study design. Eur J Heart Fail 2010 Dec;12(12):1354-1362. [doi: 10.1093/eurjhf/hfq199] [Medline: 21098580]
22. Kulshreshtha A, Kvedar J, Goyal A, Halpern E, Watson A. Use of remote monitoring to improve outcomes in patients with heart failure: a pilot trial. Int J Telemed App 2010;2010:e1-e7. [doi: 10.1155/2010/870959]
23. Nundy S, Razi RR, Dick JJ, Smith B, Mayo A, O'Connor A, et al. A text messaging intervention to improve heart failure self-management after hospital discharge in a largely African-American population: before-after study. J Med Internet Res 2013;15(3):e53 [FREE Full text] [doi: 10.2196/jmir.2317] [Medline: 23478028]
24. Scherr D, Kastner P, Kollmann A, Hallas A, Auer J, Krappinger H, et al. Effect of home-based telemonitoring using mobile phone technology on the outcome of heart failure patients after an episode of acute decompensation: randomized controlled trial. J Med Internet Res 2009;11(3):e34 [FREE Full text] [doi: 10.2196/jmir.1252] [Medline: 19687005]
25. Seto E, Leonard KJ, Cafazzo JA, Barnsley J, Masino C, Ross HJ. Mobile phone-based telemonitoring for heart failure management: a randomized controlled trial. J Med Internet Res 2012;14(1):e31 [FREE Full text] [doi: 10.2196/jmir.1909] [Medline: 22356799]
26. Zan S, Agboola S, Moore SA, Parks KA, Kvedar JC, Jethwani K. Patient engagement with a mobile web-based telemonitoring system for heart failure self-management: a pilot study. JMIR Mhealth Uhealth 2015;3(2):e33 [FREE Full text] [doi: 10.2196/mhealth.3789] [Medline: 25842282]
27. Côté I, Farris K, Feeny D. Is adherence to drug treatment correlated with health-related quality of life? Qual Life Res 2003 Sep;12(6):621-633. [Medline: 14516172]
28. Wu J, Moser DK, Chung ML, Lennie TA. Objectively measured, but not self-reported, medication adherence independently predicts event-free survival in patients with heart failure. J Card Fail 2008 Apr;14(3):203-210 [FREE Full text] [doi: 10.1016/j.cardfail.2007.11.005] [Medline: 18381183]
29. Riegel B, Knafl GJ. Electronically monitored medication adherence predicts hospitalization in heart failure patients. Patient Prefer Adherence 2013;8:1-13 [FREE Full text] [doi: 10.2147/PPA.S54520] [Medline: 24353407]
30. Riegel B, Moser DK, Anker SD, Appel LJ, Dunbar SB, Grady KL, American Heart Association Council on Cardiovascular Nursing, American Heart Association Council on Cardiovascular Nursing, American Heart Association Council on Clinical Cardiology, American Heart Association Council on Nutrition, Physical Activity,Metabolism, American Heart Association Interdisciplinary Council on Quality of CareOutcomes Research. State of the science: promoting self-care in persons with heart failure: a scientific statement from the American Heart Association. Circulation 2009 Sep 22;120(12):1141-1163 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.109.192628] [Medline: 19720935]
31. Molloy GJ, O'Carroll RE, Witham MD, McMurdo ME. Interventions to enhance adherence to medications in patients with heart failure: a systematic review. Circ Heart Fail 2012 Jan;5(1):126-133 [FREE Full text] [doi: 10.1161/CIRCHEARTFAILURE.111.964569] [Medline: 22253407]
32. Ruppar TM, Delgado JM, Temple J. Medication adherence interventions for heart failure patients: a meta-analysis. Eur J Cardiovasc Nurs 2015 Oct;14(5):395-404. [doi: 10.1177/1474515115571213] [Medline: 25669661]
33. DiMatteo MR. Variations in patients' adherence to medical recommendations: a quantitative review of 50 years of research. Med Care 2004 Mar;42(3):200-209. [Medline: 15076819]
34. van der Wal MH, van Veldhuisen DJ, Veeger NJ, Rutten FH, Jaarsma T. Compliance with non-pharmacological recommendations and outcome in heart failure patients. Eur Heart J 2010 Jun;31(12):1486-1493 [FREE Full text] [doi: 10.1093/eurheartj/ehq091] [Medline: 20436049]
35. Lainscak M, Blue L, Clark AL, Dahlström U, Dickstein K, Ekman I, et al. Self-care management of heart failure: practical recommendations from the Patient Care Committee of the Heart Failure Association of the European Society of Cardiology. Eur J Heart Fail 2011 Feb;13(2):115-126. [doi: 10.1093/eurjhf/hfq219] [Medline: 21148593]
36. Wu J, Moser DK, De Jong MJ, Rayens MK, Chung ML, Riegel B, et al. Defining an evidence-based cutpoint for medication adherence in heart failure. Am Heart J 2009 Feb;157(2):285-291 [FREE Full text] [doi: 10.1016/j.ahj.2008.10.001] [Medline: 19185635]
37. Dickstein K, Cohen-Solal A, Filippatos G, McMurray JJ, Ponikowski P, Poole-Wilson PA, ESC Committee for Practice Guidelines (CPG). ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure 2008: the Task Force for the Diagnosis and Treatment of Acute and Chronic Heart Failure 2008 of the European Society of Cardiology. Developed in collaboration with the Heart Failure Association of the ESC (HFA) and endorsed by the European Society of Intensive

Care Medicine (ESICM). Eur Heart J 2008 Oct;29(19):2388-2442 [FREE Full text] [doi: 10.1093/eurheartj/ehn309] [Medline: 18799522]
38. Sokol MC, McGuigan KA, Verbrugge RR, Epstein RS. Impact of medication adherence on hospitalization risk and healthcare cost. Med Care 2005 Jun;43(6):521-530. [Medline: 15908846]
39. Pressler SJ. Cognitive functioning and chronic heart failure: a review of the literature (2002-July 2007). J Cardiovasc Nurs 2008;23(3):239-249. [doi: 10.1097/01.JCN.0000305096.09710.ec] [Medline: 18437066]
40. Pressler SJ, Subramanian U, Kareken D, Perkins SM, Gradus-Pizlo I, Sauvé MJ, et al. Cognitive deficits in chronic heart failure. Nurs Res 2010;59(2):127-139 [FREE Full text] [doi: 10.1097/NNR.0b013e3181d1a747] [Medline: 20216015]
41. Hamine S, Gerth-Guyette E, Faulx D, Green BB, Ginsburg AS. Impact of mHealth chronic disease management on treatment adherence and patient outcomes: a systematic review. J Med Internet Res 2015;17(2):e52 [FREE Full text] [doi: 10.2196/jmir.3951] [Medline: 25803266]
42. Mistry N, Keepanasseril A, Wilczynski NL, Nieuwlaat R, Ravall M, Haynes RB, Patient Adherence Review Team. Technology-mediated interventions for enhancing medication adherence. J Am Med Inform Assoc 2015 Apr;22(e1):e177-e193 [FREE Full text] [doi: 10.1093/jamia/ocu047] [Medline: 25726568]
43. DiMatteo MR, Sherbourne CD, Hays RD, Ordway L, Kravitz RL, McGlynn EA, et al. Physicians' characteristics influence patients' adherence to medical treatment: results from the Medical Outcomes Study. Health Psychol 1993 Mar;12(2):93-102. [Medline: 8500445]
44. Sherbourne CD, Hays RD, Ordway L, DiMatteo MR, Kravitz RL. Antecedents of adherence to medical recommendations: results from the Medical Outcomes Study. J Behav Med 1992 Oct;15(5):447-468. [Medline: 1447757]
45. Kroenke K, Spitzer R. The PHQ-9: a new depression diagnostic and severity measure. Psychiat Ann 2002 Sep 01;32(9):509-515. [doi: 10.3928/0048-5713-20020901-06]
46. Pressler SJ, Subramanian U, Perkins SM, Gradus-Pizlo I, Kareken D, Kim J, et al. Measuring depressive symptoms in heart failure: validity and reliability of the patient health questionnaire-8. Am J Crit Care 2011 Mar;20(2):146-152 [FREE Full text] [doi: 10.4037/ajcc2010931] [Medline: 20378777]
47. Rector TS, Cohn JN. Assessment of patient outcome with the Minnesota Living with Heart Failure questionnaire: reliability and validity during a randomized, double-blind, placebo-controlled trial of pimobendan. Pimobendan Multicenter Research Group. Am Heart J 1992 Oct;124(4):1017-1025. [Medline: 1529875]
48. Fayers PM, Hand DJ. Factor analysis, causal indicators and quality of life. Qual Life Res 1997 Mar;6(2):139-150. [Medline: 9161114]
49. Bennet SJ, Oldridge NB, Eckert GJ, Embree JL, Browning S, Hou N, et al. Discriminant properties of commonly used quality of life measures in heart failure. Qual Life Res 2002 Jun;11(4):349-359. [Medline: 12086120]

Abbreviations

BWH: Brigham Women's Hospital
ED: emergency department
HF: heart failure
HRQoL: health-related quality of life
MEMS: medication event monitoring system
MGH: Massachusetts General Hospital
MLHFQ: Minnesota Living with Heart Failure Questionnaire
MOS: Medical Outcomes Survey
NYHA: New York Heart Association
PHQ: Patient Health Questionnaire
RCT: randomized controlled trial

Edited by G Eysenbach; submitted 29.10.15; peer-reviewed by H Nandigam, J Roettl; comments to author 26.11.15; revised version received 03.02.16; accepted 09.02.16; published 06.05.16.

Please cite as:
Hale TM, Jethwani K, Kandola MS, Saldana F, Kvedar JC
A Remote Medication Monitoring System for Chronic Heart Failure Patients to Reduce Readmissions: A Two-Arm Randomized Pilot Study
J Med Internet Res 2016;18(4):e91
URL: http://www.jmir.org/2016/4/e91/
doi:10.2196/jmir.5256
PMID:27154462

Original Paper

Mortality Benefit of a Fourth-Generation Synchronous Telehealth Program for the Management of Chronic Cardiovascular Disease: A Longitudinal Study

Chi-Sheng Hung[1], MD, PhD; Jiun-Yu Yu[2], PhD; Yen-Hung Lin[3], MD, PhD; Ying-Hsien Chen[1], MD; Ching-Chang Huang[1], MD; Jen-Kuang Lee[1], MD; Pao-Yu Chuang[4]; Yi-Lwun Ho[1], MD, PhD; Ming-Fong Chen[3], MD, PhD

[1]Telehealth Center, National Taiwan University Hospital, Taipei, Taiwan

[2]Department of Business Administration, College of Management, National Taiwan University, Taipei, Taiwan

[3]Division of Cardiology, Department of Internal Medicine, National Taiwan University Hospital, Taipei, Taiwan

[4]Department of Nursing, National Taiwan University Hospital, Taipei, Taiwan

Corresponding Author:
Yi-Lwun Ho, MD, PhD
Telehealth Center
National Taiwan University Hospital
#7, Chung-Shan South Road, Taipei, Taiwan
Taipei,
Taiwan
Phone: 886 223123456 ext 62152
Fax: 886 227044688
Email: ylho@ntu.edu.tw

Abstract

Background: We have shown that a fourth-generation telehealth program that analyzes and responds synchronously to data transferred from patients is associated with fewer hospitalizations and lower medical costs. Whether a fourth-generation telehealth program can reduce all-cause mortality has not yet been reported for patients with chronic cardiovascular disease.

Objective: We conducted a clinical epidemiology study retrospectively to determine whether a fourth-generation telehealth program can reduce all-cause mortality for patients with chronic cardiovascular disease.

Methods: We enrolled 576 patients who had joined a telehealth program and compared them with 1178 control patients. A Cox proportional hazards model was fitted to analyze the impact of risk predictors on all-cause mortality. The model adjusted for age, sex, and chronic comorbidities.

Results: There were 53 (9.3%) deaths in the telehealth group and 136 (11.54%) deaths in the control group. We found that the telehealth program violated the proportional hazards assumption by the Schoenfeld residual test. Thus, we fitted a Cox regression model with time-varying covariates. The results showed an estimated hazard ratio (HR) of 0.866 (95% CI 0.837-0.896, $P<.001$; number needed to treat at 1 year=55.6, 95% CI 43.2-75.7 based on HR of telehealth program) for the telehealth program on all-cause mortality after adjusting for age, sex, and comorbidities. The time-varying interaction term in this analysis showed that the beneficial effect of telehealth would increase over time.

Conclusions: The results suggest that our fourth-generation telehealth program is associated with less all-cause mortality compared with usual care after adjusting for chronic comorbidities.

(J Med Internet Res 2016;18(5):e102) doi:10.2196/jmir.5718

KEYWORDS

cardiovascular diseases; telemedicine; all-cause mortality; outcome assessment (health care)

Introduction

Cardiovascular disease (CVD) is a major health burden worldwide [1,2]. To improve the efficacy and reduce the burden on the health care system, telemonitoring technology has been applied to the disease management program for chronic CVD. The mortality benefit of the telehealth program for chronic heart failure has been demonstrated repeatedly [3,4]. Telemedicine

has also been shown to improve the control of vascular risk factors among patients with established cardiovascular diseases [5]. However, the evidence for the long-term (more than 1 year) benefits of a telehealth program in other chronic cardiovascular diseases is inadequate. A randomized controlled trial of a telehealth program using a nonimmediate data analysis system for 1 year resulted in increased mortality among elderly patients with chronic diseases [6]. This result raised a serious concern about the use of telehealth programs in managing patients with chronic CVD.

Because of the technologies adopted, different telehealth care programs provide different levels of care integration. Based on the integration of care, Anker and coworkers [7] classified telehealth programs into four generations. The fourth, or newest, generation of telehealth programs provides the continuous presence of a physician and nursing staff to analyze and respond synchronously to newly acquired patient data. We have shown in our prior studies that our fourth-generation telehealth care program was associated with a lower rate of emergency department visits and hospitalizations among patients with chronic CVD [8,9]. However, the long-term impact of a fourth-generation telehealth program on mortality among patients with chronic CVD has not yet been reported in the literature.

Based on our previous reports, we hypothesized that a fourth-generation telehealth program could reduce mortality in patients with chronic CVD. To test this hypothesis, we retrospectively analyzed the all-cause mortality data among 576 patients with chronic CVD who received the telehealth care program and 1175 patients who did not receive telehealth care.

Methods

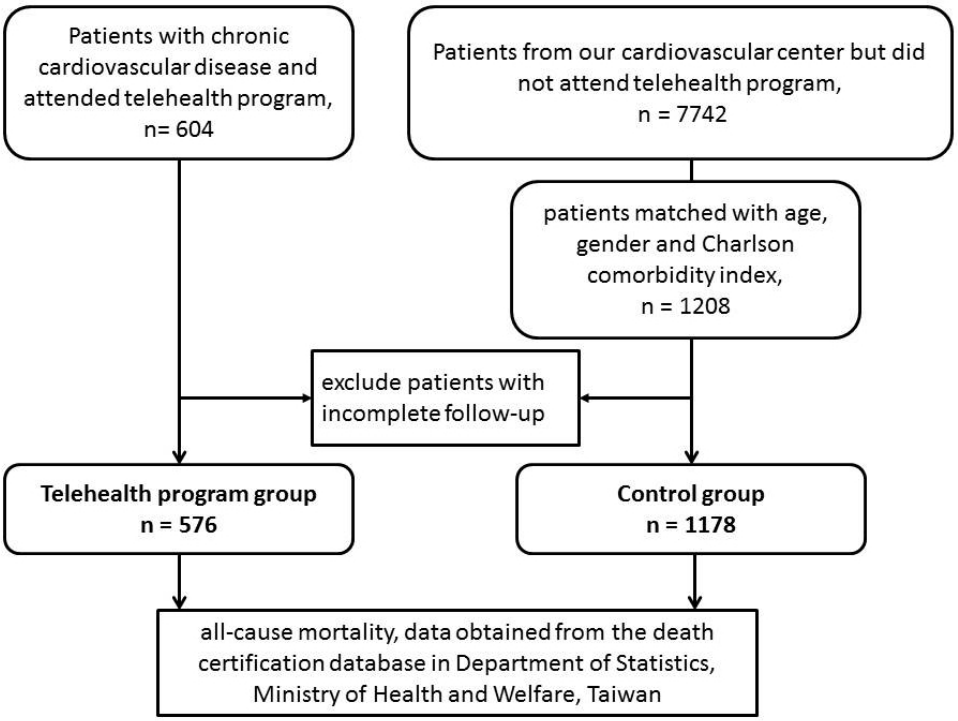
Study Design

This was a single-center, clinical retrospective epidemiologic study and was approved by the Institutional Review Board of National Taiwan University Hospital, Taipei, Taiwan. All experiments in this study were performed in accordance with relevant guidelines and regulations. Informed consent was obtained from all participants.

Recruitment

The study was conducted from December 2009 to April 2013 at the Telehealth Center of the hospital, and conducted by the Taiwan ELEctroHEALTH study group (TELEHEALTH study group). The original study method has been described previously [8]. The flow diagram of patient enrollment is shown in Figure 1. Briefly, patients older than 20 years with chronic CVD receiving the telehealth program at our telehealth center were enrolled as the study group. The decision for receiving the telehealth program was depended on patients and/or their caregivers. Chronic CVD included coronary artery disease (CAD), myocardial infarction, heart failure, peripheral artery disease (PAD), stroke, and hypertension. The control group included participants who visited our cardiovascular center during the same period, but did not participate in the telehealth care program (received usual care only). The exclusion criteria in this study (for both telehealth group and control group) included: (1) younger than 20 years, (2) patients without any one of chronic cardiovascular diseases, and (3) patients not followed in our hospital. After excluding ineligible patients, a total of 576 cases and 1178 controls were enrolled in this study.

Figure 1. The flow diagram of patient enrollment.



Telehealth Care Program

The fourth-generation telehealth program at our telehealth center is a synchronized and integrated remote management program for chronic diseases. The Internet-based platform was developed by the Graduate Institute of Biomedical Electronics and Bioinformatics, National Taiwan University, Taiwan. The details of this program have been reported previously [8]. Briefly, this telehealth program provides the following services: (1) biometric data, including single-lead electrocardiography, blood pressure, heart rate, and oximetry, are transferred from patients to our telehealth center daily and on-demand; (2) nurse case managers telephone patients daily and on-demand for communication and health promotion; (3) full-time nurse case managers and cardiologists are in charge of care 24 hours a day; and (4) long-term medication and management are discussed with the patients' primary care physician after acute events. This telehealth program bridges acute and home care and emphasizes education, prevention, and early detection of clinical deterioration.

Usual Care

Patients in the control group received the usual care provided by the primary care physicians at our cardiovascular center according to updated guidelines including, but not limited to, the American Heart Association's guidelines for lifestyle modification and primary prevention to reduce cardiovascular risk, guidelines for the management of stable ischemic heart disease, and the American Diabetes Association's guidelines for the management of diabetes. Patients made routine outpatient department visits (once every 3 months) to their primary care physicians. There was no contact between the telehealth center and patients receiving usual care.

Data Collection

All demographic and clinical data were obtained from the electronic database of the hospital. The diagnosis of a chronic disease was based on the electronic database. The discharge diagnosis was used if the outpatient and discharge diagnoses disagreed.

The primary outcome of this study was all-cause mortality. Mortality data were obtained from the death certification database in the Department of Statistics, Ministry of Health and Welfare, Taiwan.

Statistical Analysis

Statistical analysis was performed using the R version 2.14.0 software (R Foundation for Statistical Computing, Vienna, Austria). A two-sided P value ≤.05 was considered statistically significant. The continuous variables are presented as mean (SD), whereas categorical variables are presented as frequency and percentage. In univariate analysis, the potential predictive factors of all-cause mortality were examined by the chi-square test, Fisher exact test, two-sample t test, or Wilcoxon rank sum test as appropriate for the data type. Next, multivariate analysis was conducted by fitting the Cox proportional hazards model to estimate the adjusted effects of predictive factors on time to all-cause mortality.

To ensure the quality of analysis, the model-fitting techniques for (1) variable selection, (2) goodness-of-fit assessment, and (3) regression diagnostics were used in our regression analyses. Specifically, the stepwise variable selection procedure (with iterations between the forward and backward steps) was applied to obtain the candidate final Cox proportional hazards model. All the univariate significant and nonsignificant relevant covariates listed in Table 1 of Ho et al [8] (eg, age, sex, and comorbidities) and some of their interactions were put on the variable list to be selected. The significance levels for entry and for stay were set to .15 to be conservative. Simple and multiple generalized additive models (GAMs) of the binary response (1=dead vs 0=alive) were fitted to detect nonlinear effects of continuous covariates and identify appropriate cut-off point(s) for discretizing them, if necessary, during the stepwise variable selection procedure. Then, with the aid of substantive knowledge, the best candidate final Cox proportional hazards model was identified manually by dropping the covariates with P>.05 one at a time until all regression coefficients were significantly different from zero. Any discrepancy between the results of the univariate analysis and multivariate analysis was likely due to the confounding effects of uncontrolled covariates in univariate analysis or the masking effects of intermediate variables (or mediators) in the multivariate analysis.

Finally, the statistical tools of regression diagnostics for verification of proportional hazards assumption, residual analysis, detection of influential cases, and check of multicollinearity were applied to discover any model or data problems. The required proportional hazards assumption was tested based on the scaled Schoenfeld residuals. We added interaction terms between time (days) and the covariates that violated the proportional hazards assumption into the Cox proportional hazards model to examine their time-varying effects on all-cause mortality in the stepwise variable selection procedure. Technically, the original wide-form survival data were reconstructed into a long-form data structure using the so-called counting process style of input for fitting such Cox regression models. Numbers needed to treat was estimated based on the hazard ratio of telehealth program [10].

Results

Descriptive Statistics

A total of 1754 patients (576 in the telehealth group and 1178 in the control group) were enrolled in this study (Figure 1). The baseline characteristics were reported previously. Briefly, the mean age was 64.5 (SD 16.0) years and 61.17% (1073/1754) were male. At baseline, age, sex, and Charlson comorbidity index (1.35 in telehealth group vs 1.21 in the control group, P=.07) were comparable between the two groups. In the telehealth group, however, there were more patients with CAD (243/576, 42.2% vs 392/1178, 33.28% in telehealth group vs control group, respectively), heart failure (112/576, 19.4% vs 186/1178, 15.79%), stroke (71/576, 12.3% vs 110/1178, 9.34%), dementia, chronic obstructive pulmonary disease, diabetes, and peptic ulcer disease (Table 1 in [8]). The median follow-up time was 566 (IQR 349-807) days for the telehealth group and 1074 (IQR 524-1280) days for the control group. During the follow-up

period, there were less hospitalizations (mean 0.05, SD 0.12 per month vs mean 0.11, SD 0.21 per month, *P*<.001) and emergency department visits (mean 0.06, SD 0.13 per month vs mean 0.09, SD 0.23 per month, *P*<.001) in the telehealth group compared with the control group. The outpatient visit times were comparable between the two study groups (mean 1.57, SD 1.12 per month vs mean 1.66, SD 1.78 per month, *P*=.75).

Survival Outcome

There were 53/576 (9.3%) deaths in the telehealth group and 136/1178 (11.50%) deaths in the control group. The Kaplan-Meier survival curve is shown in Figure 2. The estimated survival curves of the two study groups were similar during the follow-up period (log rank test: *P*=.81). Because the baseline comorbidities were heterogeneous and the follow-up durations differed between the two study groups, Cox proportional hazards analyses were performed. A proportional hazards assumption was tested for each variable using the Schoenfeld residuals test. The result showed the proportional hazards assumption was violated for the telehealth program (Table 1). The time-varying telehealth program effect was then applied to the Cox regression model to examine its effect on mortality. The time-varying Cox regression analysis revealed a hazard ratio (HR) of 0.866 (95% CI 0.837-0.896, *P*<.001) for the telehealth program on all-cause mortality. The time-dependent interaction term in this Cox regression analysis indicated that the beneficial effect of the telehealth program on the HR would increase over time. Thus, the estimated numbers needed to treat for the telehealth program to prevent one death at 1, 2, and 3 years were 55.6, 40.5, and 27.7, respectively (Table 2).

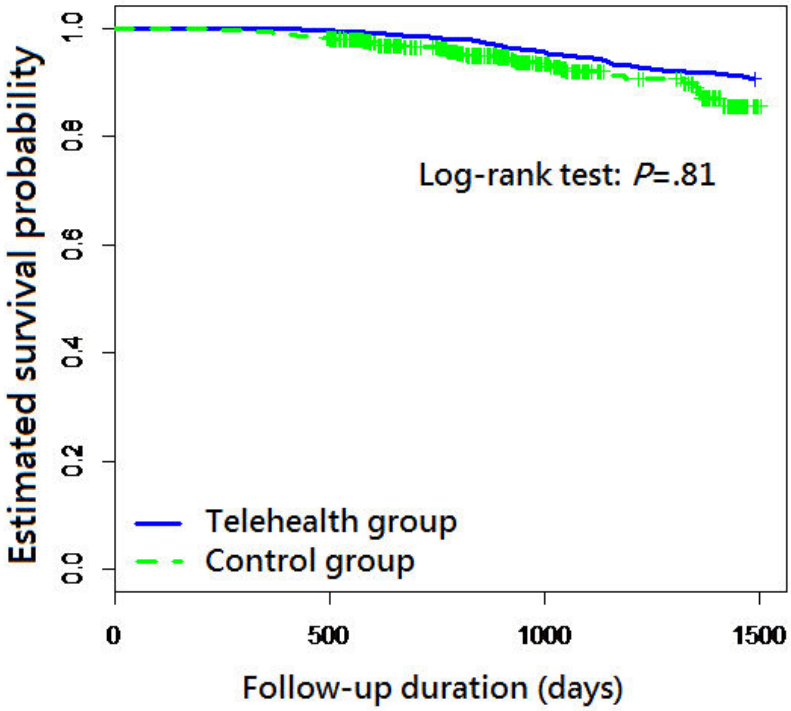
Table 1. The estimated adjusted hazard ratios for the clinical predictors of all-cause mortality obtained from a multivariate Cox regression model.

Clinical predictors	HR (95% CI)	*P*
Age (years)	1.019 (1.018-1.021)	<.001
Age > 69.809 (years)	1.890 (1.810-1.974)	<.001
Age > 69.809 (years) × telehealth program	0.837 (0.788-0.889)	<.001
Male	1.152 (1.125-1.179)	<.001
Telehealth program	0.866 (0.810-0.926)	<.001
Telehealth program × time-to-mortality (days)	0.9997 (0.9996-0.9998)	<.001
Heart failure	1.941 (1.891-1.992)	<.001
Myocardial infarction	1.097 (1.051-1.146)	<.001
Coronary artery disease	0.843 (0.823-0.865)	<.001
Peripheral arterial disease	0.824 (0.786-0.864)	<.001
Dementia	1.204 (1.115-1.301)	<.001
Chronic obstructive pulmonary disease	1.147 (1.107-1.189)	<.001
Connective tissue disease	1.949 (1.784-2.130)	<.001
Liver disease, mild	1.263 (1.187-1.344)	<.001
Liver disease, moderate to severe	1.876 (1.663-2.117)	<.001
Diabetes mellitus	1.291 (1.259-1.325)	<.001
Diabetes mellitus with end organ damage	0.775 (0.713-0.843)	<.001
Chronic kidney disease, moderate to severe	2.362 (2.295-2.430)	<.001
Malignancy	3.403 (3.313-3.496)	<.001
Malignancy with metastasis	2.816 (2.644-3.000)	<.001
Peptic ulcer disease	1.260 (1.196-1.328)	<.001

Table 2. The estimated numbers needed to treat at various times after telehealth program (based on the hazard ratio of the telehealth program).

Time from treatment	Number needed to treat	95% CI
1 year	55.6	43.2-75.7
2 years	40.5	31.8-54.7
3 years	27.7	22.0-37.0

Figure 2. The Kaplan-Meier estimate of the survival curves for all-cause mortality.



Discussion

Principal Results

The major finding of this study is that our fourth-generation telehealth program was associated with lower all-cause mortality after controlling for multiple comorbidities among patients with chronic cardiovascular disease. Prior studies on the effects of telehealth programs on all-cause mortality among patients with chronic diseases revealed mixed results [11-13]. This study demonstrated the clear benefits of a newer generation of telehealth care among a population with a broader range of cardiovascular diseases during a longer follow-up duration. In addition, our prior study demonstrated that a fourth-generation telehealth program was independently associated with lower numbers of emergency department visits, hospitalizations, and medical costs (US $587.60 vs US $1163.60 per month in the telehealth and control groups, respectively, $P<.001$) [8]. Based on this result and on our previous reports, our data show a fourth-generation telehealth program is cost-effective and lifesaving.

Comparison With Previous Work

Among patients with chronic heart failure, the beneficial effect of telehealth care on all-cause mortality has been repeatedly demonstrated in studies [3,4]. The effect of telehealth care among patients with a broader range of chronic comorbidities, however, is still controversial. In the Whole System Demonstrator (WSD) study including 3230 participants with diabetes, chronic obstructive pulmonary disease, or heart failure, a second-generation telehealth program was related to lower all-cause mortality [12]. In a recent study conducted by the Mayo Clinic on 205 frail, older adults with a broader range of chronic diseases, a second generation of telemonitoring resulted in higher mortality compared with usual care [6]. Several factors may contribute to the contradictory results between these two studies and our own studies.

First, the telehealth program has changed over time. Early telemonitoring systems used asynchronous techniques to record and process the collected subjective symptoms and physiological data. A response to the data collected could take more than 1 day during the off hours. This newer system uses a synchronous technique to respond to the information received to act in a timely manner. Through quick communication, more accurate diagnoses and decisions can be made. Second, the age of the study participants differs. The mean ages in the studies were 80.3 years in the Mayo study, 70.3 years in the WSD study, and 64.5 years in our study. The potential barriers for elderly participants to effectively use a telehealth program include a lack of skill to operate a new device, psychological resistance to new technology, and the presence of more comorbidities. Third, the comorbidity profiles also differed. The reported chronic comorbidities were as follows: a mean Charlson score of 2.9 in the Mayo study, a mean number of chronic conditions of 1.8 in the WSD study, and a mean Charlson score of 1.3 in our study. In a meta-analysis on the effectiveness of telemonitoring for four chronic conditions, the results were more consistent for cardiac and pulmonary conditions compared with diabetes and hypertension [14]. Whether or not telehealth programs are effective for all types of chronic conditions is still unknown. The difference in the telemonitoring technology, age, and comorbidities may contribute to the different results from the three studies. These differences warrant further investigation in the future.

Our study found that the effect of the telehealth program on mortality was time dependent; namely, that the beneficial effect

increased with longer follow-up durations. Although an increasing benefit on mortality has not previously been reported, an increasing benefit on other surrogate endpoints has been noted previously. In a study of telemedicine conducted with elderly Medicare patients, glycated hemoglobin (HbA1c) and blood pressure levels were lower after the follow-up for 5 years. The difference between the telemedicine and control groups increased further at 5 years [15]. It is plausible that the continuous improvement in the control of chronic conditions, including hypertension and hyperglycemia, gradually led to a benefit in mortality. Moreover, continuous education via daily telephone communication could have improved the knowledge and techniques used by the caregivers. Through repeated discussion of the condition, caregivers learn to better manage the acute exacerbation. This may also have contributed to further improvements in outcomes during the follow-up period.

The paradoxical associations of CAD and PAD with lower all-cause mortality found in our data were unexpected. Plausible explanations of these survival paradoxes included chance or unmeasured residual confounding factors, especially effective treatment for CAD and PAD. In a recent large cohort of 102,023 patients with stable CAD, a history of percutaneous coronary intervention or coronary artery bypass graft in the past 6 months were associated with hazard ratios of 0.651 and 0.516, respectively, on all-cause mortality [16]. It is possible that our results reflect the unmeasured treatment effect of these two diseases. In addition, the patients in our telehealth group had higher rates of CAD and PAD compared with those in the control group. Although we had evaluated the collinearity between different predictors with the variance inflation factor, it is still possible that the paradoxical protective effect reflected the protective effect of the telehealth program, rather than the effect of CAD or PAD.

Limitations

Our study has some limitations. First, because our study was not a randomized study, unmeasured confounding factors might have influenced the results. Second, because our study compared the whole service program and not any single monitoring device or physiological parameter, the effect of each device or parameter on mortality cannot be assessed separately. Third, because all-cause mortality instead of mortality from specific causes was used in this study, we cannot draw a definite conclusion regarding the effects of telehealth monitoring on specific causes of death. Finally, the numbers needed to treat is estimated based on the hazard ratio of the telehealth program. This may lead to overestimating the numbers needed to treat.

Conclusions

Our data suggest that a fourth-generation telehealth program is independently associated with lower all-cause mortality among patients with chronic CVD and multiple comorbidities.

Acknowledgments

The Taiwan National Science Council (grant no. NSC 102-2220-E-002-008), the National Taiwan University Hospital (grant no. 105-S3139) and the National Science Council Center for Dynamical Biomarkers and Translational Medicine, National Central University, Taiwan (grant no. NSC 101-2911-I-008-001).

Authors' Contributions

PYC, YLH, and MFC designed the study; YHL, YHC, JKL, and PYC collected the data; CSH, JYY, YHC, CCH, JYY, and YHL analyzed the data; and CSH, JYY, and YLH wrote the manuscript.

Conflicts of Interest

None declared.

References

1. Kung H, Hoyert DL, Xu J, Murphy SL. Deaths: final data for 2005. Natl Vital Stat Rep 2008 Apr 24;56(10):1-120 [FREE Full text] [Medline: 18512336]
2. GBD 2013 Mortality Causes of Death Collaborators. Global, regional, and national age-sex specific all-cause and cause-specific mortality for 240 causes of death, 1990-2013: a systematic analysis for the Global Burden of Disease Study 2013. Lancet 2015 Jan 10;385(9963):117-171. [doi: 10.1016/S0140-6736(14)61682-2] [Medline: 25530442]
3. 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]
4. Inglis SC, Clark RA, McAlister FA, Stewart S, Cleland JG. Which components of heart failure programmes are effective? A systematic review and meta-analysis of the outcomes of structured telephone support or telemonitoring as the primary component of chronic heart failure management in 8323 patients: Abridged Cochrane Review. Eur J Heart Fail 2011 Sep;13(9):1028-1040. [doi: 10.1093/eurjhf/hfr039] [Medline: 21733889]
5. Vernooij JW, Kaasjager HA, van der Graaf Y, Wierdsma J, Grandjean HM, Hovens MM, SMARTStudy Group. Internet based vascular risk factor management for patients with clinically manifest vascular disease: randomised controlled trial. BMJ 2012;344:e3750 [FREE Full text] [Medline: 22692651]
6. Takahashi PY, Shah ND, Naessens JM. Telehealth monitoring with nurse clinician oversight-reply. Arch Intern Med 2012 Nov 12;172(20):1612-1613. [doi: 10.1001/2013.jamainternmed.276] [Medline: 23752871]

7. Anker SD, Koehler F, Abraham WT. Telemedicine and remote management of patients with heart failure. Lancet 2011 Aug 20;378(9792):731-739. [doi: 10.1016/S0140-6736(11)61229-4] [Medline: 21856487]
8. Ho Y, Yu J, Lin Y, Chen Y, Huang C, Hsu T, et al. Assessment of the cost-effectiveness and clinical outcomes of a fourth-generation synchronous telehealth program for the management of chronic cardiovascular disease. J Med Internet Res 2014;16(6):e145 [FREE Full text] [doi: 10.2196/jmir.3346] [Medline: 24915187]
9. Chen Y, Lin Y, Hung C, Huang C, Yeih D, Chuang P, et al. Clinical outcome and cost-effectiveness of a synchronous telehealth service for seniors and nonseniors with cardiovascular diseases: quasi-experimental study. J Med Internet Res 2013;15(4):e87 [FREE Full text] [doi: 10.2196/jmir.2091] [Medline: 23615318]
10. Altman DG, Andersen PK. Calculating the number needed to treat for trials where the outcome is time to an event. BMJ 1999 Dec 4;319(7223):1492-1495 [FREE Full text] [Medline: 10582940]
11. Takahashi PY, Pecina JL, Upatising B, Chaudhry R, Shah ND, Van HH, et al. A randomized controlled trial of telemonitoring in older adults with multiple health issues to prevent hospitalizations and emergency department visits. Arch Intern Med 2012 May 28;172(10):773-779 [FREE Full text] [doi: 10.1001/archinternmed.2012.256] [Medline: 22507696]
12. Steventon A, Bardsley M, Billings J, Dixon J, Doll H, Hirani S, Whole System Demonstrator Evaluation Team. Effect of telehealth on use of secondary care and mortality: findings from the Whole System Demonstrator cluster randomised trial. BMJ 2012;344:e3874 [FREE Full text] [Medline: 22723612]
13. Bashshur RL, Shannon GW, Smith BR, Alverson DC, Antoniotti N, Barsan WG, et al. The empirical foundations of telemedicine interventions for chronic disease management. Telemed J E Health 2014 Sep;20(9):769-800 [FREE Full text] [doi: 10.1089/tmj.2014.9981] [Medline: 24968105]
14. Paré G, Jaana M, Sicotte C. Systematic review of home telemonitoring for chronic diseases: the evidence base. J Am Med Inform Assoc 2007;14(3):269-277 [FREE Full text] [doi: 10.1197/jamia.M2270] [Medline: 17329725]
15. Shea S, Weinstock RS, Teresi JA, Palmas W, Starren J, Cimino JJ, IDEATel Consortium. A randomized trial comparing telemedicine case management with usual care in older, ethnically diverse, medically underserved patients with diabetes mellitus: 5 year results of the IDEATel study. J Am Med Inform Assoc 2009;16(4):446-456 [FREE Full text] [doi: 10.1197/jamia.M3157] [Medline: 19390093]
16. Rapsomaniki E, Shah A, Perel P, Denaxas S, George J, Nicholas O, et al. Prognostic models for stable coronary artery disease based on electronic health record cohort of 102 023 patients. Eur Heart J 2014 Apr;35(13):844-852 [FREE Full text] [doi: 10.1093/eurheartj/eht533] [Medline: 24353280]

Abbreviations

CAD: coronary artery disease
CVD: cardiovascular disease
GAM: generalized additive models
HR: hazard ratio
PAD: peripheral artery disease
WSD: Whole System Demonstrator

Edited by G Eysenbach; submitted 03.03.16; peer-reviewed by E Krupinski, P Yellowlees, J Pecina; comments to author 28.03.16; revised version received 11.04.16; accepted 12.04.16; published 13.05.16.

Please cite as:
Hung CS, Yu JY, Lin YH, Chen YH, Huang CC, Lee JK, Chuang PY, Ho YL, Chen MF
Mortality Benefit of a Fourth-Generation Synchronous Telehealth Program for the Management of Chronic Cardiovascular Disease: A Longitudinal Study
J Med Internet Res 2016;18(5):e102
URL: http://www.jmir.org/2016/5/e102/
doi:10.2196/jmir.5718
PMID:27177497

Original Paper

Exploring Concordance of Patient-Reported Information on PatientsLikeMe and Medical Claims Data at the Patient Level

Gabriel S Eichler[1], PhD; Elisenda Cochin[1], B A; Jian Han[2], PhD; Sylvia Hu[2], PhD; Timothy E Vaughan[1], PhD; Paul Wicks[1], PhD; Charles Barr[2], MD, MPH; Jenny Devenport[2], PhD

[1]PatientsLikeMe, Cambridge, MA, United States

[2]Genentech, Medical Affairs, South San Francisco, CA, United States

Corresponding Author:
Paul Wicks, PhD
PatientsLikeMe
160 Second Street
Cambridge, MA, 02142
United States
Phone: 1 617 314 4352
Fax: 1 8668506240
Email: pwicks@patientslikeme.com

Related Article:

This is a corrected version. See correction statement: http://www.jmir.org/2016/10/e271

Abstract

Background: With the emergence of data generated by patient-powered research networks, it is informative to characterize their correspondence with health care system-generated data.

Objectives: This study explored the linking of 2 disparate sources of real-world data: patient-reported data from a patient-powered research network (PatientsLikeMe) and insurance claims.

Methods: Active patients within the PatientsLikeMe community, residing in the United States, aged 18 years or older, with a self-reported diagnosis of multiple sclerosis or Parkinson's disease (PD) were invited to participate during a 2-week period in December 2014. Patient-reported data were anonymously matched and compared to IMS Health medical and pharmacy claims data with dates of service between December 2009 and December 2014. Patient-level match (identity), diagnosis, and usage of disease-modifying therapies (DMTs) were compared between data sources.

Results: Among 603 consenting patients, 94% had at least 1 record in the IMS Health dataset; of these, there was 93% agreement rate for multiple sclerosis diagnosis. Concordance on the use of any treatment was 59%, and agreement on reports of specific treatment usage (within an imputed 5-year period) ranged from 73.5% to 100%.

Conclusions: It is possible to match patient identities between the 2 data sources, and the high concordance at multiple levels suggests that the matching process was accurate. Likewise, the high degree of concordance suggests that these patients were able to accurately self-report their diagnosis and, to a lesser degree, their treatment usage. Further studies of linked data types are warranted to evaluate the use of enriched datasets to generate novel insights.

(J Med Internet Res 2016;18(5):e110) doi:10.2196/jmir.5130

KEYWORDS

feasibility of data linking; patient-powered research network

Introduction

Health researchers today have the luxury of choosing from a multitude of "big data" sources. Given the growing diversity of data, new insights and transformative potential arise when complementary sources can be linked at the individual level (with appropriate consents and protections) to provide a more holistic view of a patient's journey with illness or to inform medical decisions [1].

Although the potential advantages of data linkage and integration of existing sources are tantalizing (eg, rapid, economical access), rigorous scrutiny is required to assure the resultant data, and analyses provide meaningful and valid insights. Key questions and potential hurdles must be addressed before data integration can be performed reliably at scale. Comparisons between data from clinical encounters (captured in electronic medical records, claims, or administrative records) and patient-reported data (captured via telephone interviews, in person interviews [2], mailed surveys, and in-clinic surveys [3,4]) have been undertaken in previous research. It has been recognized that there is a need to confirm data reported by patients engaged through Web-based methods. Cascade et al used manual medical chart review to validate data from 50 patients with gout registered on MediGuard.org and were able to verify diagnosis in 76% of patients who consented to participate [5]. However, paper record extraction was a barrier to validation with physicians' offices being slow to return charts or, in some cases, refusing to take part altogether. The concordance of patient-reported data collected by an Internet-based registry with data captured from clinical care has not been well characterized in the literature.

In this paper, we describe the evaluation of the potential benefits and challenges of linking data from an online patient community (PatientsLikeMe [PLM]) with administrative claims data. Electronic medical and pharmacy claims data are collected as a result of clinical care, primarily for billing and reimbursement purposes. In addition, they are widely used by a variety of stakeholders for other purposes, including practice management, identifying market trends, studying medication compliance, conducting outcomes research, and building health economic models. However, such secondary uses of billing data may result in limitations when used for research purposes. For example, it may be difficult to infer causality because transactional data report *what* happened without *why* it happened. Furthermore, changes in billing data may reflect administrative errors or limitations in classifications rather than actual changes to patient condition, evaluations completed, and treatments given.

PLM data have been used in over 65 peer-reviewed scientific publications, particularly in the areas of patient-centered outcomes research [6], development of new patient-reported outcomes [7], and clinical trials [8]. As a patient-powered research network (PPRN) covering over 2000 diseases, the system emphasizes patient-reported data submitted in the Web and allows patients to enter detailed pseudoanonymous records of their illness using structured and quantitative tracking tools. However, PLM has limitations because, currently, all data are patient-reported and no validating information is required. Thus, some members may not be formally diagnosed with the conditions they report and some could intentionally (or unintentionally) enter erroneous data. The quality of self-reported data may be related to patients' level of technological comfort, cognitive abilities, degree of motivation, and willingness to return to the site to enter data longitudinally. Websites dependent on voluntary self-reporting suffer from significant attrition over time [9]. As a newer form of data in the health care ecosystem, there are additional steps to be performed such as comparing data quality to traditional data and checking for external validity that might affect generalizability of findings [8].

Objectives

In this pilot study, we investigated the feasibility of linking a small sample of patients from 2 large data sources at the patient level: the IMS Health database of medical and retail pharmacy claims, covering >250 million lives in the United States, and PLM, a PPRN hosting over 325,000 patients with chronic life-changing illnesses [6]. A secondary objective was to describe the concordance of some common data elements between the 2 datasets.

Methods

Data Sources

Members of PLM are prompted to voluntarily report information on their condition history (eg, first symptom date and diagnosis date), relevant symptoms, treatments taken, and laboratory results, and broader health metrics such as weight and quality of life. For the purposes of this pilot study, diagnosis status and treatments were selected as the entities most suitable for matching with the IMS Health database.

Medical and pharmacy billing data are available for purchase for the purposes of research. For this study, an extract of medical and pharmacy (preadjudicated) claims from the IMS Integrated Data Warehouse with dates of service from December 2009 to December 2014 was used. Claims were from office-based physicians and specialists (noncash visits only). Pharmacy claims included prescription data collected from retail, long-term care, specialty, and mail order computerized pharmacy records (both insured and cash transactions). Although the Integrated Data Warehouse data source is not a closed system (ie, not all records from all patients are captured), coverage is wide, representing more than 1 billion medical service records annually as well as an estimated 75% of all prescriptions dispensed in the United States.

Patients

Multiple sclerosis (MS) was the condition of primary interest in this study, PD patients were included as non-MS controls as a comparator to evaluate the accuracy of MS diagnosis reporting. From the range of common conditions on the PLM system at the time of study, both conditions had relatively large engaged communities, were diagnosed by specialists, have relatively low misdiagnosis rates, and can be identified by the presence of DMTs that are quite specific to each condition. Based on prior studies conducted by PLM, and the pilot nature of this study, target enrollment for this study was 350 actively engaged patients from the PLM communities for MS and PD (approximately 4:1 ratio based on prevalence of the reported diagnoses in the PLM database).

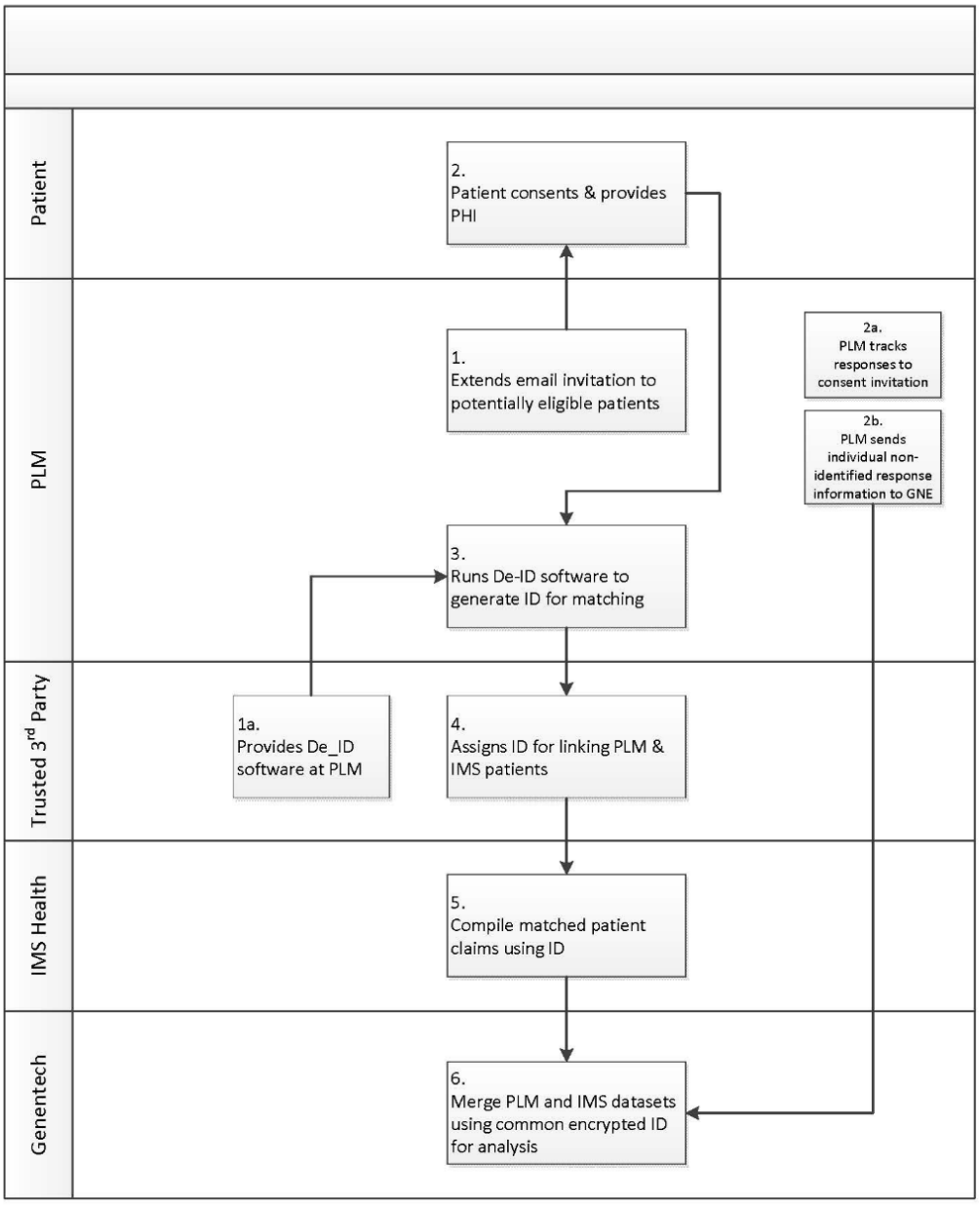
The recruiting process occurred over a 2-week period in early December 2014 and included PPRN patients who met the following eligibility criteria: (1) reported MS or PD on their profile; (2) "actively engaged," defined as having logged into the PPRN at least once in the 90 days before December 8, 2014;

(3) aged 18 years or older at the time of study; (4) account was maintained by the patient; and (5) resided in the United States in the last 10 years.

Eligible patients were sent an invitation on December 8, 2014 to participate in the study via email and private message. Patients who reviewed the message within the 2-week enrollment period could click on a link to a research subject information page where a written statement of research information with informed consent was presented. The research protocol was approved by the New England Institutional Review Board on December 3, 2014. Eligible consenting participants were asked to provide their name, previous last names, date of birth, and zip codes for the last 10 years. This personal health information provided during consent was entered into third-party software to generate encrypted, deidentified (De-ID) tokens. The tokens were then used to make a deterministic match to similarly anonymized patients with claims in the IMS Health database. No payments were made to patients either for taking part in this study or for joining PLM.

No data were exchanged between the 2 dataset owners, and all personal health information remained at each source. Data from consenting patients were sent directly to Genentech (GNE) for analysis using encrypted IDs, generated from third party software, for the match. A detailed review of the linking methodology is beyond the scope of this paper. However, extensive research has previously been conducted using this methodology [10], and Figure 1 provides an overview of the matching process used here.

Figure 1. Overview of record matching process.



Analysis Methods

Detailed patient disposition summaries were completed. Descriptive statistics of demographic characteristics were calculated for invited eligible patients, consenting patients (with or without a match), and nonconsenting patients. Additional evaluations of PPRN profile data were explored for patients without matching identity, diagnosis, or drug information in the claims dataset.

The primary objective of the analysis was to explore the feasibility of patient-level linking in the 2 datasets. The overall patient match rate between the 2 sources was calculated as the percentage of consenting patients from PLM with a matching record (De-ID) in the claims database. The secondary objectives were to further explore the concordance between the data sources at the diagnosis and, for MS patients, treatment levels.

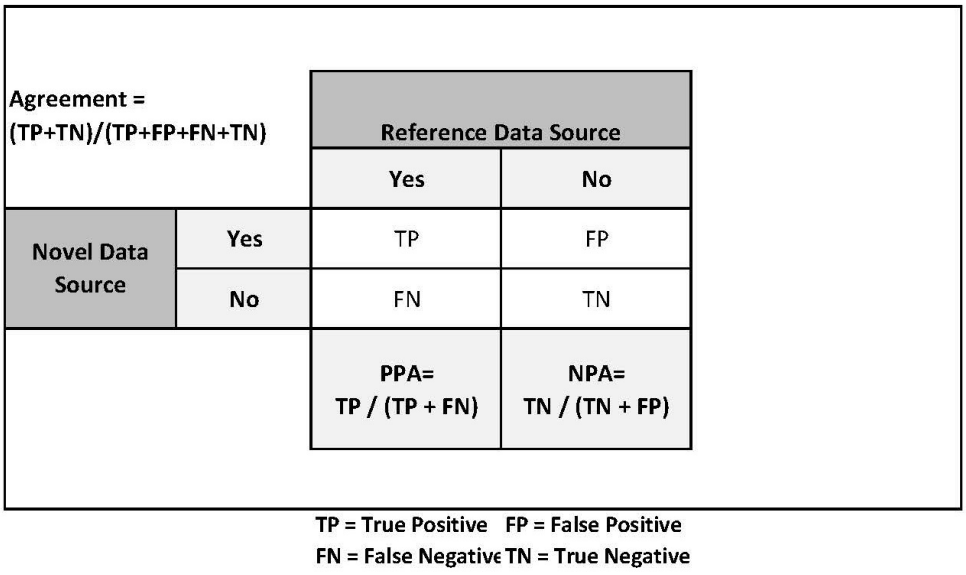
Two-by-two tables were constructed to evaluate the concordance of diagnosis (MS or not MS) and DMT use (none vs any), and overall (raw) agreement was calculated. Positive percent agreement (PPA) and negative percent agreement (NPA) were calculated, using IMS claims data as the reference (Figure 2). Because the extracted claims dataset may not include all the claims that patients have and treatment reporting in PLM is completely optional, neither data source can truly be considered the "gold standard" and so some discordance between the 2 data sources was expected.

For the evaluation of diagnoses (MS vs not MS), a patient was considered to have MS in the PLM dataset if they had reported MS as a condition in their profile. In the event that a patient reported both MS and PD in their profile, the patient was classified as MS. A patient was considered to have MS in the IMS dataset if they either had at least one medical claim with the International Classification of Diseases, 9th Revision (ICD9) diagnosis code for MS ("340") or they had a pharmacy claim for one of the following known MS medications, brand name (chemical name, manufacturer): Ampyra (dalfampridine, Acorda), Aubagio (teriflunomide, Genzyme), Avonex (interferon beta 1-a, Biogen), Betaseron (interferon beta 1-b, Bayer), Copaxone (glatiramer acetate, Teva), Extavia (interferon beta-1b, Novartis), Gilenya (fingolimod, Novartis), Lemtrada (alemtuzumab, Genzyme), Novantrone (mitoxantrone, EMD-Serono), Rebif (interferon beta 1-a, EMD-Serono & Pfizer), Tecfidera (dimethyl fumarate, Biogen), Tysabri (natalizumab, Biogen), and Plegridy (peginterferon beta-1a, Biogen).

For the evaluation of MS treatments, the analysis population was restricted to patients with MS as identified in their PLM profile. Analyses were completed for all data (any DMT vs no DMT) and for specific DMTs (yes or no). DMTs include all those drugs listed previously except Ampyra (dalfampridine), which is indicated for the improvement of walking in MS. Spelling variations for DMTs in both datasets were manually adjudicated to ensure appropriate matching. Additional analyses were completed limiting treatment data to those patients reporting treatment within the last 5 years, corresponding to the claims data available for this project. Because many patients did not report start or stop dates for their medications, some analyses were also completed with missing dates imputed as the system-generated dates on which the patients entered specific treatments in their profiles. This imputation permitted additional patient treatment records to fall within the last 5 years.

All analyses were completed using SAS software, version 9.2 for PC (Copyright © 2002-2008, SAS Institute Inc.). SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc., Cary, NC, USA.

Figure 2. Calculation of agreement.



Results

Of the 5364 patients invited via private message through the PPRN website, 2039 (38.0%) viewed the initial invitation, 3325 (61.9%) patients did not open the invitation and, therefore, did not actively accept or decline to participate in the study. Of the 2039 patients who viewed the study invitation, 1421 patients failed to complete the questionnaire, 15 patients actively declined to consent to take part in the study, and, ultimately, 603 patients consented to the study and completed the questionnaire (30% participation rate [603 of 2039] and 11% response rate [603 of 5364]). The target recruitment goal was reached in 24 hours, though the questionnaire remained available for 15 days. Among the 603 consenting patients, 414 reported having MS (69%), 188 reported having PD (31%), and 1 reported having both MS and PD.

Patient-Level Match

Overall, 567 of the 603 consenting patients (94%) were found in the claims dataset (based on a record with a corresponding De-ID token) and were, therefore, revealed to be real patients with confirmed identities for the purposes of this study. A total of 36 patients were not found within the claims dataset. Two patients subsequently asked to be excluded from the analysis and the PLM community, so the final linked dataset contained 565 matched patients.

Demographics

Demographic characteristics from PLM profiles for those patients who were invited, consented, matched, and did not match can be found in Table 1. Patients who consented had higher rates of PD, diagnosis by physician (for either condition), reported DMT use, and nonmissing insurance type than nonconsenting patients. All of these factors, except PD diagnosis, would be expected to be related to likelihood of match because diagnosis and treatment by a physician would increase the likelihood of medical and/or pharmaceutical services within claims data.

Concordance of Diagnosis

The patient-reported MS diagnosis status was confirmed in the claims dataset for 524 (92.7%) matched patients (Table 2). PPA—having MS diagnosis in PLM when MS diagnosis exists in IMS Health—was 99.7%; whereas NPA—not having MS reported in PLM when no MS claim exists in IMS—was 81.1%.

Further analyses were undertaken to better understand potential reasons for mismatch on diagnosis. The 41 patients with discordant MS status in the 2 data sources tended to have fewer medical and pharmacy claims, mean (standard deviation) values are 16.6 (20.74) and 92.9 (242.10), respectively, than patients with concordant MS status, mean (standard deviation) values are 43.1 (44.40) and 206.9 (236.70), respectively. Although the sample sizes were small, patients with concordant MS status were more likely to have a physician diagnosis (96%) than patients with discordant status (88%). Discordant and concordant patients tended to have similar diagnosis date distributions (ie, duration of illness) as well as similar types of insurance (eg, missing, private employer or union based, government funded).

Table 1. Demographic characteristics.

Patient Characteristics[a]	Patients invited[b] (N=5,362)	Patients who did not consent (N=4,759)	Consenting patients with claims[b] match (N=565)	Consenting patients with no claims match (N=36)
Age in yrs (SD)	54.7 (11.53)	54.3 (11.60)	57.4 (10.63)	56.5 (11.02)
Number of Females (%)	3,546 (66.1)	3,136 (65.9)	385 (68.1)	25 (69.4)
Primary condition in PLM				
MS (%)	3,869 (72.2)	3,470 (72.9)	379 (67.1)	20 (55.6)
PD (%)	1,333 (24.9)	1,151 (24.2)	168 (29.7)	14 (38.9)
Other (%)	160 (2.9)	138 (2.9)	18 (3.2)	2 (5.6)
Patient Reports MS or PD Diagnosed by Physician (%)	4,512 (84.2)	3,934 (73.4)	544 (96.3)	34 (94.4)
PLM patients with MS as primary or secondary condition (%)	3,976 (74.2)	3,564 (74.9)	392 (69.4)	21 (58.3)
MS subtype (% MS)				
Relapsing-Remitting	2,429 (61.1)	2,165 (60.8)	250 (63.8)	14 (66.7)
Primary progressive	253 (6.4)	227 (6.4)	25 (6.4)	1 (4.76)
Secondary progressive	551 (13.9)	470 (13.2)	78 (19.9)	3 (14.3)
Progressive relapsing	127 (3.2)	116 (3.3)	10 (2.6)	1 (4.8)
Unreported	616 (15.5)	586 (16.5)	29 (7.4)	2 (9.5)
Years since MS Diagnosis[a] (%)				
0 - ≤5 Years	581 (14.6)	524 (14.7)	53 (13.5)	4 (19.05)
>5 - ≤10 Years	1,102 (27.7)	980 (27.5)	116 (29.6)	6 (28.6)
>10 - ≤15 Years	711 (17.9)	619 (17.4)	87 (22.2)	5 (23.8)
>15 - ≤20 Years	408 (10.3)	358 (10.0)	48 (12.2)	2 (9.5)
>20 Years	566 (14.3)	492 (10.8)	71 (18.1)	3 (14.3)
[Not Reported]	609 (15.3)	591 (16.6)	17 (4.3)	1 (4.8)
Reported MS DMT use in PLM (%)	3,118 (78.4)	2,751 (77.2)	351 (90.0)	16 (76.2)
Reported Insurance Type (%)				
Indian Health Service	1 (0.02)	1 (0.02)	0 (0.0)	0 (0.0)
Medicaid/ other low-income plan	195 (3.64)	165 (3.47)	29 (5.13)	1 (2.78)
Medicare	1023 (19.08)	799 (16.79)	209 (36.99)	15 (41.67)
National health service	7 (0.13)	7 (0.15)		
Other type of insurance	49 (0.91)	42 (0.88)	6 (1.06)	1 (2.78)
Private (individual plan)	210 (3.92)	183 (3.85)	27 (4.78)	
Private (via employer /union)	1351 (25.20)	1141 (23.98)	203 (35.93)	7 (19.44)
TRICARE (or oth military ins)	55 (1.03)	47 (0.99)	7 (1.24)	1 (2.78)
Veteran's Administration	69 (1.29)	54 (1.13)	13 (2.30)	2 (5.56)
No Insurance	87 (1.62)	79 (1.66)	6 (1.06)	2 (5.56)
Prefer not to answer	77 (1.44)	74 (1.55)	2 (0.35)	1 (2.78)
[Not Reported]	2238 (41.74)	2167 (45.53)	63 (11.15)	6 (16.67)

[a]Source for all characteristics is PLM; all statistics reported are n (%) unless otherwise noted.

[b]Two patients who were invited, consented and had at least 1 claim in the claims dataset asked to have their profiles removed from PLM and are, therefore, not represented in this analysis.

Table 2. Concordance of data sources on diagnosis.

MS reported by patient[a]	MS claim[b]			Agreement[c]		
	MS	Not MS	Total	Overall	PPA	NPA
MS	352	40	392	92.7%	99.7%	81.1%
Not MS	1	172	173			
Total	353	212	565			

[a]MS reported by patient means that patient did or did not report MS in their patient profile.

[b]MS claim means that patient did or did not have a claim with a diagnosis code for MS or a claim for a drug uniquely indicated for treatment of MS.

[c]PPA and NPA calculations use claims datasource as reference.

Concordance of DMT Use for Patients Reporting MS

Patients reporting MS on their profile had an overall agreement of 58.7% for DMT usage between the datasets, with high PPA (97%) and low NPA (18.3%; top section of Table 3). To explore the possibility that low overall agreement was due to higher rates of DMT use more than 5 years ago (before claims extract for this analysis), patient-reported DMT use was categorized based on use within the past 5 years only. For the purposes of determining patient-reported medication use within 5 years, missing DMT use dates were treated as no DMT within 5 years and, separately, imputed using date of medication data entry by patient (second and third sections of Table 3). A shift of patients to "no DMT use within 5 years" in the PPRN source was observed, but overall agreement between data sources remained similar. When missing dates were imputed, the number of patients using DMTs within 5 years increased from 188 to 284, PPA was increased by 25.4%, and NPA was decreased by 23.6%.

Analyses of specific DMT usage based on 5-year categories with imputation of missing dates revealed an overall agreement ranging from 73.5% to 100% depending on the DMT. PPA and NPA were >50% for most DMTs. Complete results are provided in Table 4.

Table 3. Concordance of data sources on any DMT use in MS patients.[a]

Imputation	Patient-reported MS medication	MS medication in claims (5 years)			Agreement[b]		
		Any DMT	No DMT	Total	Overall	PPA	NPA
	Overall any DMT	195	156	351	58.7%	97.0%	18.3%
No	Overall no DMT	6	35	41			
	Total	201	191	392			
No	Any DMT (5 years)	111	77	188	57.4%	55.2%	59.7%
	No DMT (5 years)	90	114	204			
	Total	201	191	392			
Yes	Any DMT (5 years)	162	122	284	58.9%	80.6%	36.1%
	No DMT (5 years)	39	69	108			
	Total	201	191	392			

[a]DMT use categories in this table reflect claims or patient reported; if no dates available then included under "no" for no imputation; missing dates for DMT use imputed based on date of medication entry to evaluate use within 5 years in section marked "yes."

[b]PPA and NPA calculations use claims as reference.

Table 4. Concordance of data sources on specific DMT use within 5 years in PLM MS patients.[a]

	Patient-reported MS medication	MS medication in claims		Agreement[b]		
		Yes	No	Overall	PPA	NPA
Alemtuzumab	Yes	0	0	100.0%	-	100.0%
	No	0	392			
Dimethyl fumarate	Yes	47	60	83.4%	90.4%	82.4%
	No	5	280			
Fingolimod	Yes	24	27	91.8%	82.8%	92.6%
	No	5	336			
Glatiramer acetate	Yes	49	72	74.0%	62.0%	77.0%
	No	30	241			
Interferon Beta 1a	Yes	37	86	73.5%	67.3%	74.5%
	No	18	251			
Interferon Beta 1b	Yes	11	38	87.2%	47.8%	89.7%
	No	12	331			
Metoxantrone	Yes	0	17	95.7%	-	95.7%
	No	0	375			
Natalizumab	Yes	15	51	84.2%	57.7%	86.1%
	No	11	315			
Peginterferon Beta 1	Yes	0	0	99.7%	0.0%	100.0%
	No	1	391			
Teriflunomide	Yes	3	13	94.9%	30.0%	96.6%
	No	7	369			

[a]DMT use categories in this table reflect claims or patient reported; missing dates for DMT use imputed based on date of medication entry to evaluate use within 5 years.

[b]PPA and NPA calculations use IMS claims as reference.

Discussion

Principal Findings

This study demonstrated the feasibility of linking patient-reported data with billing claims data generated in the clinical setting. The study surpassed its expected degree of patient engagement (n=350), having gained consent from 603 patients. This suggests that patients were more open to, and supportive of, this project than first anticipated.

The degree of concordance between the PPRN and claims datasets was high, with 94% agreement on patient identity, 93% agreement on MS diagnosis status (MS vs not MS), and agreement on specific DMT usage ranged from 74% to 100%. Finally, the results indicate that these patients were willing and able to accurately recount their diagnosis and to a lesser degree, their use of DMTs in MS care. In total, these findings support the conclusion that accurate linkages at the patient level are possible, opening the doors for further research on an enriched dataset.

Although the rates of patient matching and concordance of diagnosis were high, a careful review of findings and limitations is important to better understand the context and implications for future research. For example, some discordance or nonoverlap of the 2 data sources should be expected a priori. At the patient level, corresponding record(s) in the data sources would not be expected in the following situations: if patients paid for physician services in cash, were uninsured, insured by payers not reflected in the IMS Health database, or received medications via distributers not reflected in the IMS Health database. Analysis of PLM patient profile data revealed that the 36 patients without any claims had lower rates of private insurance than the 565 patients with claims (19.4% vs 35.9%, respectively). Thus, it is possible that at least some of these patients had care that was not captured in the claims dataset.

At the diagnosis and treatment levels, a match to the claims system would not be expected in several situations such as if a patient received treatment for their condition before December 2009 or too recently (because there is a lag between date of service and billing claim submission) or joined a disease community without having been diagnosed by a physician. Conversely, we might see claims without corresponding information in the PPRN if the patient forgot or chose not to report certain information. Patients reported both current and retrospective diagnoses and treatments—so timing did not appear to be a major driver of match rates. Patients with a match

were more likely to report having a physician diagnosis (96%) than patients without a match (88%). A larger percentage of diagnosis-unmatched patients reported having Primary Progressive MS, than those who did have a matching diagnosis (15% vs 4%, respectively). Because there are no approved treatments for this subtype of MS, these patients are less likely to be taking a DMT and are thus less likely to have a treatment claim in the IMS database [11,12]. Finally, patients without matching diagnoses in both datasets tended to have fewer medical and pharmacy claims than patients with a match, thus decreasing the chances of a match on a diagnosis category or specific treatment.

The characteristics of each dataset must also be considered to better understand the potential value of linking them to create an enriched dataset. The PLM community is a self-selected population that is biased somewhat to be more educated, female, white, and technologically savvy [13,14]. PLM collects voluntary reports that may result in data gaps on care received or health status. In addition, the patients who consented into this study may represent a more engaged and activated population than those individuals who did not consent. Although this study does represent the findings of linking the PLM PPRN to a claims dataset, attempting to generalize the findings to other studies of linked data sources without regard to their designs, patient inclusion factors, and response characteristics may be unrealistic.

Medical and prescription claims databases are a widely used tool for exploring how real-world healthcare services and treatments are used by physicians and patients. Claims data has long been a core resource for health economics and outcomes research, therapeutic persistence and utilization studies, quality, and other core areas of health services research [15]. Nonetheless, claims data has potentially significant limitations including the time lag of 6-9 months in complete reporting because of the claims filing and adjudication processes [15,16], the inability to capture clinical outcomes of health care interventions more precisely, and the lack of explanations for why something happened. For example, when an MS patient experiences a relapse in their disease they may present to a hospital, be subjected to multiple tests, be provided new medications, and even switched to another DMT. Although the services and treatments used may be captured within the claims dataset, the initiating event, the relapse, is never explicitly captured in the claims dataset, forcing the health services researchers to infer the relapse event using a complex and likely imperfect algorithm [15,17]. The potential real-time nature of patient-generated and patient-reported data provide an opportunity to capture an entirely new dimension of the patient's health care experience, that is, their account of what happened, although these are of course subject to self-reporting biases [18].

Although the levels of concordance and generalizability of these results to diseases other than MS has not been demonstrated within this study, the high degree of concordance between the patient-reported and claims-based datasets observed here suggests that it may be possible to use these types of combined datasets to answer new research questions. For example, because the PPRN captures treatment experience data (side effect severity, specific side effects, perceived efficacy, and so forth) one could create treatment persistence curves that are stratified by the patient's experience. This would enable researchers to get at the "why" behind treatment decision starts and stops. Similarly, given that PLM enables patients to synchronize wearable devices to their profile [19], one could look at how health care utilization patterns may differ based on the physical activity of a patient. In addition, patient preferences might be obtained from questionnaires and then linked to their claims data to quantify differences in health care utilization. These sorts of questions would be difficult or impossible to answer in the absence of a linked dataset.

Conclusion

Real-world datasets and patient-reported data are becoming increasingly powerful tools for research, quality improvement, and broader understanding of the evolving health care system. The further integration of traditional claims or healthcare-system generated data, coupled with patient-reported data will continue to bring about ever new opportunities to study disease, patient experience, and the applications of health care services and their effects on the health care system and patient outcomes. The generalizability of our findings and patient willingness to consent to data linkage for clinical research needs further exploration in other settings.

Acknowledgments

The authors would like to acknowledge and thank the contributing patients on PLM who participated in this project. In addition, Jamie Heywood, Ben Heywood, Pronabesh DasMahapatra, Jennifer Kissinger, Bill Evans, Erica Chain, and David Purdie for their roles in helping the team prepare for and execute this project. IMS Health, Inc. provided the prescription and medical claims for the consenting PLM patients in this study, coordinated the de-ID engine process, and provided an aggregate report (not contained in this manuscript). The team is especially grateful for the expert consultation of Jason Lucas and Thomas Haskell from IMS Health, Inc. throughout the deidentification and matching processes.

Authors' Contributions

The study was co-conceived by all the authors. All the authors contributed to the study design. GE, EC, TV, and PW completed data collection. The analysis was led by JD, JH, and SH. All the authors participated in the interpretation of the data. GE, EC, JD, and CB led on drafting the manuscript. All the other authors contributed to drafting the manuscript. All the authors have read and approved the final manuscript.

Conflicts of Interest

CEB, SH, JH, and JD are employed by Genentech. EC, PW, TV, and GE are employees of PLM. PLM receives funding from Abbvie, Accorda, Actelion, Amgen, AstraZeneca, Avanir, Biogen, Boehringer Ingelheim, Genzyme, Janssen, Johnson & Johnson, Merck, Novartis, Sanofi, and UCB. The PLM R&D team has received research grant funding from Kaiser Permanente, the Robert Wood Johnson Foundation, Sage Bionetworks, the AKU Society, and the University of Maryland. This work was supported by Genentech.

References

1. Weber GM, Mandl KD, Kohane IS. Finding the missing link for big biomedical data. JAMA 2014 Jun 25;311(24):2479-2480. [doi: 10.1001/jama.2014.4228] [Medline: 24854141]
2. Corser W, Sikorskii A, Olomu A, Stommel M, Proden C, Holmes-Rovner M. Concordance between comorbidity data from patient self-report interviews and medical record documentation. BMC Health Services Research 2008 Apr 16;8:85 [FREE Full text] [doi: 10.1186/1472-6963-8-85]
3. Fowles J, Fowler E, Craft C. Validation of claims diagnoses and self-reported conditions compared with medical records for selected chronic disease. Journal of Ambulatory Care Management 1998;21:1.
4. Skinner K, Miller D, Lincoln E, Lee A, Kazis L. Concordance between respondent self-reports and medical records for chronic conditions: experience from the Veterans Health Study. J Ambul Care Manage 2005;28(2):102-110. [Medline: 15923944]
5. Cascade E, Marr P, Winslow M, Burgess A, Nixon M. Conducting research on the Internet: medical record data integration with patient-reported outcomes. J Med Internet Res 2012;14(5):e137 [FREE Full text] [doi: 10.2196/jmir.2202] [Medline: 23060427]
6. Lavallee DC, Wicks P, Alfonso CR, Mullins CD. Stakeholder engagement in patient-centered outcomes research: high-touch or high-tech? Expert Rev Pharmacoecon Outcomes Res 2014 Jun;14(3):335-344. [doi: 10.1586/14737167.2014.901890] [Medline: 24661181]
7. Rothman M, Gnanaskathy A, Wicks P, Papadopoulos EJ. Can we use social media to support content validity of patient-reported outcome instruments in medical product development? Value Health 2015 Jan;18(1):1-4 [FREE Full text] [doi: 10.1016/j.jval.2014.10.001] [Medline: 25595228]
8. Wicks P, Vaughan T, Heywood J. Subjects no more: what happens when trial participants realize they hold the power? BMJ 2014;348:g368 [FREE Full text] [Medline: 24472779]
9. 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]
10. IMS Health. Real-World Evidence & HEOR Bibliography URL: http://www.imsheorbibliography.com/ [accessed 2016-04-28] [WebCite Cache ID 6h7HVSpLn]
11. Goodman A. National Multiple Sclerosis Society. 2009. Talking About Primary-Progressive MS (PPMS) URL: http://www.nationalmssociety.org/NationalMSSociety/media/MSNationalFiles/Brochures/Brochure-Primary-Progressive-MS.pdf [accessed 2016-04-28] [WebCite Cache ID 6h7HbZqy8]
12. Fitzner D, Simons M. Chronic progressive multiple sclerosis - pathogenesis of neurodegeneration and therapeutic strategies. Curr Neuropharmacol 2010 Sep;8(3):305-315 [FREE Full text] [doi: 10.2174/157015910792246218] [Medline: 21358979]
13. Bove R, Secor E, Healy BC, Musallam A, Vaughan T, Glanz BI, De Jager Philip L. Evaluation of an online platform for multiple sclerosis research: patient description, validation of severity scale, and exploration of BMI effects on disease course. PLoS One 2013;8(3):e59707 [FREE Full text] [doi: 10.1371/journal.pone.0059707] [Medline: 23527256]
14. Dimova S, De La Loge C, Durgin T, Mueller K, Lafosse C, Massagli M, et al. Characteristics of Users of the Epilepsy Community of PatientsLikeMe®: An Update. 2011. URL: http://patientslikeme_posters.s3.amazonaws.com/2012_Characteristics%20of%20Epilepsy%20Community%20of%20PLM:%20Update.pdf [accessed 2016-04-28] [WebCite Cache ID 6h7IofHp4]
15. Crystal S, Akincigil A, Bilder S, Walkup JT. Studying prescription drug use and outcomes with medicaid claims data: strengths, limitations, and strategies. Med Care 2007 Oct;45(10 Supl 2):S58-S65 [FREE Full text] [doi: 10.1097/MLR.0b013e31805371bf] [Medline: 17909385]
16. Mass Technology Leadership Council. acceleratedcure.org. 2013. Big Data and Life Sciences URL: http://www.acceleratedcure.org/sites/default/files/images/Life_Sciences_Report_Final_WEB.pdf> [accessed 2016-04-28] [WebCite Cache ID 6h7IvXJ5k]
17. Chubak J, Pocobelli G, Weiss NS. Tradeoffs between accuracy measures for electronic health care data algorithms. J Clin Epidemiol 2012 Mar;65(3):343-349.e2 [FREE Full text] [doi: 10.1016/j.jclinepi.2011.09.002] [Medline: 22197520]
18. Cook C. Mode of administration bias. J Man Manip Ther 2010 Jun;18(2):61-63 [FREE Full text] [doi: 10.1179/106698110X12640740712617] [Medline: 21655386]
19. McIninch J, Datta S, DasMahapatra P. Remote Tracking of Walking Activity in MS Patients in a Real-world Setting. In: AAN. 2015 Presented at: American Academy of Neurology Annual Meeting; 2015; Washington, DC URL: http://www.neurology.org/content/84/14_Supplement/P3.209

Edited by G Eysenbach; submitted 14.09.15; peer-reviewed by B Shen, R Haase, K Doing-Harris, P Jongen, B Craig, J Parkin Kullmann; comments to author 17.10.15; revised version received 25.11.15; accepted 21.02.16; published 12.05.16.

Please cite as:
Eichler GS, Cochin E, Han J, Hu S, Vaughan TE, Wicks P, Barr C, Devenport J
Exploring Concordance of Patient-Reported Information on PatientsLikeMe and Medical Claims Data at the Patient Level
J Med Internet Res 2016;18(5):e110
URL: http://www.jmir.org/2016/5/e110/
doi:10.2196/jmir.5130
PMID:27174602

Original Paper

A New Cuffless Device for Measuring Blood Pressure: A Real-Life Validation Study

Tessa S Schoot[1], MSc; Mariska Weenk[2], MD; Tom H van de Belt[3], PhD; Lucien JLPG Engelen[3]; Harry van Goor[2], MD, PhD, FRCS; Sebastian JH Bredie[4], MD, PhD

[1]Radboud University Medical Center, Department of Internal Medicine, Radboud University, Nijmegen, Netherlands

[2]Radboud University Medical Center, Department of Surgery, Radboud University, Nijmegen, Netherlands

[3]Radboud REshape Innovation Center, Radboud University Medical Center, Radboud University, Nijmegen, Netherlands

[4]Radboud University Medical Center, Department of Internal Medicine and Radboud REshape Innovation Center, Radboud University, Nijmegen, Netherlands

Corresponding Author:
Sebastian JH Bredie, MD, PhD
Radboud University Medical Center
Department of Internal Medicine and Radboud REshape Innovation Center
Radboud University
PO Box 9101
Nijmegen, 6533 AR
Netherlands
Phone: 31 243618819
Fax: 31 246365126
Email: bas.bredie@radboudumc.nl

Related Articles:

Comment in: http://www.jmir.org/2018/10/e10089/

Comment in: http://www.jmir.org/2018/10/e12477/

Abstract

Background: Cuffless blood pressure (BP) monitoring devices, based on pulse transit time, are being developed as an easy-to-use, more convenient, fast, and relatively cheap alternative to conventional BP measuring devices based on cuff occlusion. Thereby they may provide a great alternative to BP self-measurement.

Objective: The objective of our study was to evaluate the performance of the first release of the Checkme Health Monitor (Viatom Technology), a cuffless BP monitor, in a real-life setting. Furthermore, we wanted to investigate whether the posture of the volunteer and the position of the device relative to the heart level would influence its outcomes.

Methods: Study volunteers fell into 3 BP ranges: high (>160 mmHg), normal (130–160 mmHg), and low (<130 mmHg). All requirements for test environment, observer qualification, volunteer recruitment, and BP measurements were met according to the European Society of Hypertension International Protocol (ESH-IP) for the validation of BP measurement devices. After calibrating the Checkme device, we measured systolic BP with Checkme and a validated, oscillometric reference BP monitor (RM). Measurements were performed in randomized order both in supine and in sitting position, and with Checkme at and above heart level.

Results: We recruited 52 volunteers, of whom we excluded 15 (12 due to calibration failure with Checkme, 3 due to a variety of reasons). The remaining 37 volunteers were divided into low (n=14), medium (n=13), and high (n=10) BP ranges. There were 18 men and 19 women, with a mean age of 54.1 (SD 14.5) years, and mean recruitment systolic BP of 141.7 (SD 24.7) mmHg. BP results obtained by RM and Checkme correlated well. In the supine position, the difference between the RM and Checkme was >5 mmHg in 17 of 37 volunteers (46%), of whom 9 of 37 (24%) had a difference >10 mmHg and 5 of 37 (14%) had a difference >15 mmHg.

Conclusions: BP obtained with Checkme correlated well with RM BP, particularly in the position (supine) in which the device was calibrated. These preliminary results are promising for conducting further research on cuffless BP measurement in the clinical and outpatient settings.

(J Med Internet Res 2016;18(5):e85) doi:10.2196/jmir.5414

KEYWORDS

hypertension; cuffless blood pressure monitor; wearable device; cardiovascular risk management; patient empowerment

Introduction

Noninvasive blood pressure (BP) monitors based on cuff occlusion are used widely in and outside of care facilities. These devices measure systolic (SBP) and diastolic blood pressure (DBP) by auscultation [1] or oscillometry [2]. Disadvantages of these measurements are discomfort for the patient because of painful cuff inflation, which may influence BP outcome, and the impossibility of continuous or semicontinuous BP monitoring due to the necessity of cuff inflation and deflation. Measurements can also vary between users, for example, patients or health care workers, due to interindividual differences in use. Although self-measurement of BP using noninvasive BP monitors has been shown to produce significantly greater BP reduction in patients with hypertension than standard care using clinic-based BP measurements [3], it is not common practice because it is time consuming and has high overall costs because of expensive equipment and technologies [4].

To overcome the disadvantages of BP measurements based on cuff occlusion and to provide easy-to-use devices for reliable self-measurement, pocket-sized BP monitoring devices without the need of a pressure cuff have been developed and are entering the consumer market. The majority of the cuffless devices indirectly measure BP by determining pulse transit time, the time interval required for a pressure wave in the arterial tree to travel between 2 sites (ie, a proximal and a distal point). Pulse transit time is closely related to BP via arterial compliance. Not only are these devices able to measure BP quickly and conveniently, but some of them also measure other modalities such as pulse rate, oxygenation, respiratory rate, and skin temperature. Furthermore, with respect to BP measurement, correct cuff size and cuff position are no longer important issues to take into account for obtaining reliable results. Altogether, these new cuffless devices could be an excellent alternative to BP measuring devices based on cuff occlusion, especially for the purpose of self-measurement.

The Checkme Health Monitor (Viatom Technology, Shenzen, People's Republic of China) is a newly released Conformité Européenne-approved cuffless BP monitoring device. Checkme is a IIa category medical device compliant with directive 93/42/European Economic Community. As it is aimed at the consumer market, it has been defined as a screening device for primary medical checking and not for diagnostic use. However, for its use in a clinical setting, especially during monitoring of hypertension treatment, the device's accuracy in persons with BPs outside the normal range has to be determined as well.

To ensure the accuracy of new BP monitoring devices, several protocols have been established, such as the European Society of Hypertension International Protocol (ESH-IP) revision 2010 [5] and protocols of the Association for the Advancement of Medical Instrumentation [6,7]. However, a single unified protocol for all types of BP monitoring devices is still under development. For example, the ESH-IP and Association for the Advancement of Medical Instrumentation protocols stipulate the use of a mercury sphygmomanometer as the reference device, whereas the International Organization for Standardization protocol allows use of any type of reference manometer, as long as it meets the accuracy requirement. Furthermore, the protocols that have been developed for validating noninvasive BP devices are designed primarily for monitors that are intrinsically able to give absolute BP readings in a single measurement.

Other category devices, such as Checkme, require patient-specific calibration by a secondary measurement method or device before they can give absolute BP readings. A protocol for validating such a monitor must include provisions to assess the monitor's accuracy in tracking intrapatient BP changes, relative to the calibrated level, after a patient-specific calibration or between calibrations [8].

Another issue in daily practice is that oscillometric devices for the noninvasive estimation of BP have progressively become the clinical standard because of the need to train staff in determining BP by auscultation, cost, and the banning of mercury in many states and countries [2]. Therefore, it is conceivable that new devices are being evaluated in comparison with the easy-to-use automated oscillometric BP devices used in daily practice.

Finally, with classic BP devices, a correct BP can only be determined with the detection point (eg, the arm) at heart level. Because of the assumed method of BP measuring with cuffless devices, it is still unclear whether the device's position relative to the heart may influence the results of the measurement.

The aim of this study was to evaluate the performance of the first release of the Checkme cuffless BP monitor in a real-life patient setting. To this purpose, we compared Checkme BP measurements with measurements from a validated oscillometric reference BP monitor (RM) according to ESH-IP requirements. Our second aim was to investigate whether the posture of the volunteer and the position of the device relative to the heart level would affect outcomes.

Methods

Checkme

Checkme is a cuffless BP monitoring device, which only determines SBP. It can be used both in clinical settings and for self-measurement (Figure 1).

This biometrical device can also measure skin temperature, heart rate, oxygen saturation, and 1-lead electrocardiogram, and it can be used as a sleep monitor. Before being able to measure SBP with Checkme, a personal profile containing sex, age, weight, and height has to be created, and the device has to be calibrated with an RM. Calibration is performed by simultaneously measuring SBP with Checkme and with RM and entering the SBP of the RM into Checkme after each measurement. After both calibration measurements, the Checkme is ready to use. SBP, heart rate, and oxygen saturation can then be measured by putting the right index finger beneath the lid on top, the right thumb on the metal plate on front, and the right middle finger on the metal plate on the back. Simultaneously, the metal plate on the left side of the device has to be pressed against the palm of the left hand (Figure 2,Figure 3).

Checkme has to be held still at heart level during a measurement. Performing one measurement takes about 20 seconds. To evaluate the result, data can be transferred via Bluetooth to a mobile phone or tablet (supported operating systems are iOS or Android) with the Checkme app (Figure 4). Details by which the Checkme measures BP have not been described in the public domain.

Figure 1. Checkme Health Monitor (Viatom Technology) device.

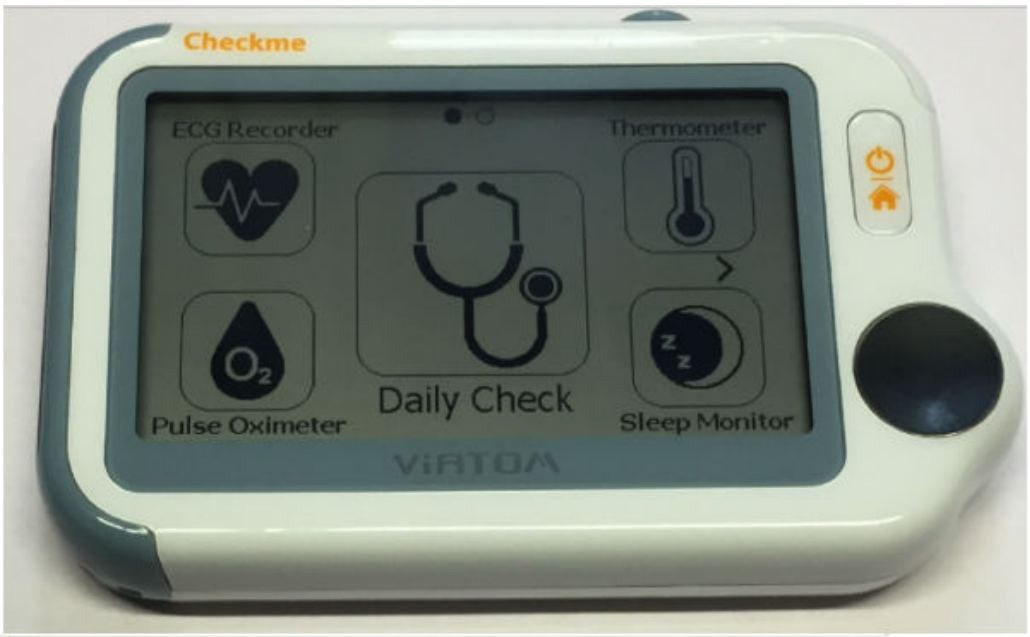


Figure 2. Checkme position during measurement (front).

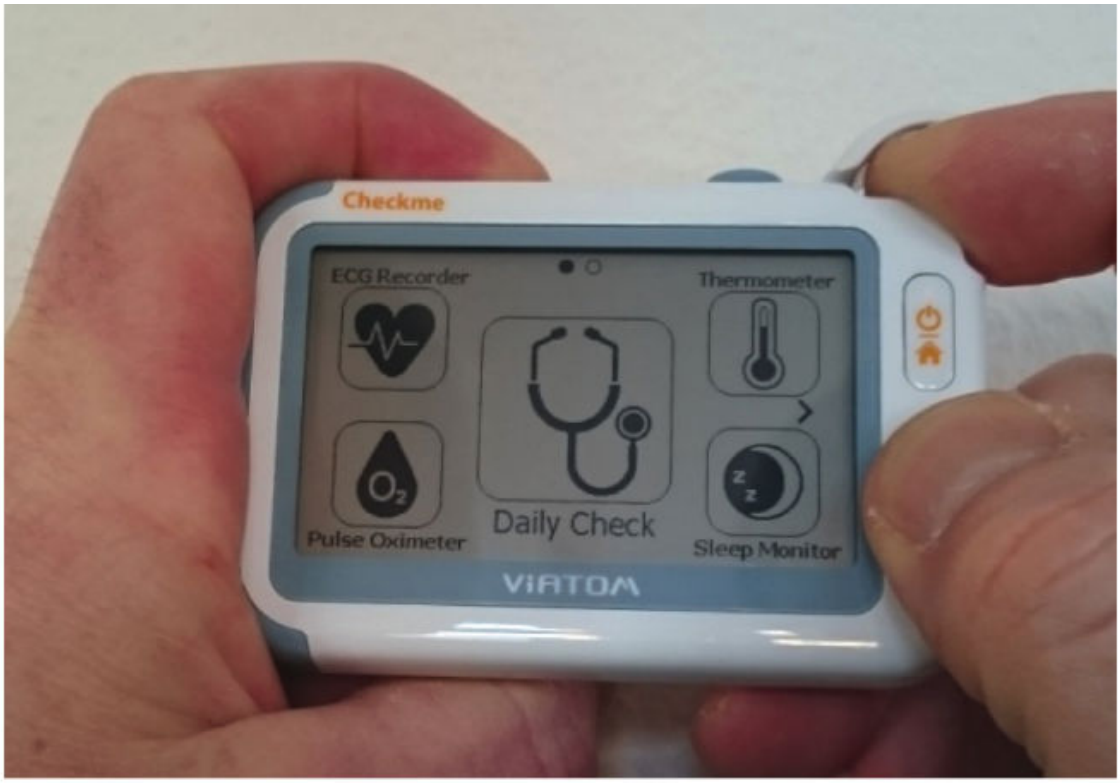


Figure 3. Checkme position during measurement (back).

Reference Device

We used the validated Vital Signs Monitor 300 series (Welch Allyn, Skaneateles Falls, NY, USA) as RM. This automatic device measures SBP and DBP in the upper arm by oscillometry. The normal adult cuff size is suitable for people with an arm circumference of 25.3–34.4 cm. We used the small adult cuff when arm circumference was lower (range 20.0–27.0 cm) and the large adult cuff when arm circumference was higher (range 40.7–55.0 cm).

Familiarization

Before the validation procedure, we took a multiple series of test measurements using the Checkme and RM to familiarize ourselves with the devices. To test the study procedure and familiarize ourselves with it, we measured 2 volunteers accordingly. We encountered no problems. Experienced technicians of the Radboud University Medical Center maintained and calibrated the RM according to the manufacturer's protocol.

Recruitment

We recruited study volunteers from patients who visited the hypertension outpatient services of the Radboud University Medical Center Department of Internal Medicine. To cover inclusion in all BP categories in this study, we also recruited patients with hypertension admitted to the hospital (highest BP range) and healthy employees (lowest BP range). We stopped recruitment after obtaining valid measurements of 37 volunteers with baseline BP measurements in the required ranges. Exclusion criteria were cardiac arrhythmias, upper-arm circumference outside the cuff range, and age <25 years. Information on age, sex, and use of antihypertension medication was collected and height, body weight, and arm circumference were measured. All volunteers gave written informed consent. The institutional review board gave permission for this study (Medical Research Ethics Committee CMO no. 2015-1717).

Protocol

This study followed the ESH-IP requirements for test environment, observer qualification, volunteer recruitment, and BP measurements for the validation of BP measurement devices [5]. Because device readings are digital, 1 researcher performed all measurements. In addition to the ESH-IP requirements, we took measurements in different positions to establish the influence of posture on device readings.

Figure 4. The Checkme app, showing heart rate (HR), electrocardiogram (EGC), and oxygen saturation (SpO_2).

Each volunteer was seen individually in a quiet, temperature-controlled room. Appropriate cuff size (in the case of RM) was chosen based on upper-arm circumference. For each individual volunteer, a new profile was created on the Checkme device, with input of sex, date of birth, height, and weight. Volunteers were given oral instructions regarding proper use of the Checkme device before measurements were taken.

Baseline measurements were performed with the volunteer in the supine position after resting for 10 minutes. BP was measured 3 times at the right upper arm with the RM. The mean of the last 2 values was used as the baseline value, on the basis of which volunteers were divided into 1 of 3 BP categories: high (SBP >160 mmHg), normal (SBP ≥130 and ≤160 mmHg), or low (SBP <130 mmHg) BP, according to ESH-IP, with at least 10 volunteers in each BP category.

Next, we calibrated the Checkme device with the volunteer in the supine position with hands resting on the lower abdomen. The last measured baseline SBP with the RM was used as the input value for calibration. After calibration, we randomized the order of measurements. In the first series of measurements, BP was measured in the supine position with Checkme at heart level (arms resting on lower abdomen), Checkme above heart level (arms stretched above the head at a 90° angle with the body), and the RM (right upper arm) according to the randomization order. After the first series of measurements in the supine position, volunteers were asked to sit up. After 5 minutes of rest, the volunteer's BP was again measured in random order with the Checkme at heart level and RM, both in the upright position. All of the above measurements were executed 3 times successively. According to ESH-IP, the interval between consecutive measurements was between 30 and 60 seconds. Failed measurements were repeated up to a maximum of 3 times.

Statistical Analysis

All statistic calculations were performed with IBM SPSS version 20 (IBM Corporation). To evaluate the influence of the volunteer's position on the device readings, we compared the means of 3 consecutive measurements with a device in the supine or sitting position by paired samples *t* -test. A difference with $P<.05$ was considered to be significant.

Results

We excluded 15 of 52 recruited volunteers: 12 due to repeated BP calibration failures with Checkme, 2 because they appeared to have low BP (SBP <130 mmHg) with already sufficient data, and 1 who declined to continue after inclusion. None of the volunteers had arrhythmias. In <3% of all measurements, BP had to be measured again due to failure during the first attempt of both the RM and the Checkme readings.

Study Population

Of the 37 volunteers who completed the study, 14 were in the low range (SBP <130 mmHg), 13 were in the medium range (SBP between 130 and 160 mmHg), and 10 were in the high range (SBP >160 mmHg). Table 1 shows their baseline characteristics. There were 18 men and 19 women with a mean age of 54.1 (SD 14.5) years. The mean baseline SBP was 141.7 (SD 24.7) mmHg. For 31 of the 37 volunteers (84%) we used the normal cuff size of the RM. Due to an arm circumference above than normal range, the remaining 6 volunteers (16%) required the large cuff.

Feasibility

In 22 of 52 volunteers (42%), calibration with Checkme failed the first time (error message: "unstable measure, calibration failed"). We repeated the procedure up to a maximum of 5 times. In 5 of 52 volunteers (10%), calibration succeeded after the second attempt, in 4 (8%) after the third attempt, and in 1 (2%) after the fifth attempt. Calibration continued to fail in 12 of 52 volunteers (23%), whereupon they were excluded from further measurements. In 2 of 37 volunteers who completed the study, the SBP measurement could not be determined in the upright position.

Comparing BP Results (Primary Aim)

Table 2 shows the BP results for RM and Checkme. Table 3 shows the proportion of patients with differences between RM and Checkme of >5, >10, and >mmHg. We constructed Bland-Altman scatter plots of BP differences between RM and Checkme against the mean BP of the RM and Checkme in the supine (Figure 5) and upright positions (Figure 6). BP results correlated with the position of Checkme relative to the heart level.

Table 1. Study population characteristics.

Characteristics	All volunteers (n=37)
Male:female	18:19
Age in years, mean (SD)	54.1 (14.5)
White, n (%)	36 (97)
Black, n (%)	1 (3)
Height in m, mean (SD)	172.2 (7.5)
Weight in kg, mean (SD)	83.3 (18.4)
Use of blood pressure-lowering drugs, n (%)	22 (60)
Normal cuff size, n (%)	31 (84)
Baseline systolic blood pressure in mmHg, mean (SD)	141.7 (24.7)

Table 2. Systolic blood pressure measurements (mmHg) taken by the reference monitor and Checkme in the supine and upright positions.

Volunteers' position	Mean[a]	SD	Range (min; max)[a]
Supine position			
Reference monitor	136.6	21.8	84.7 (106.3; 191.0)
Checkme at heart level	138.4	25.2	94.5 (94.5; 189.0)
Checkme above heart	130.7[b]	27.7	101.0 (86.0; 187.0)
Upright position			
Reference monitor	139.2	22.3	100.7 (102.3; 203.0)
Checkme at heart level	136.6[c]	25.9	87.7 (102.3; 190.0)

[a]The average or range of 3 consecutive blood pressure measurements.

[b]*P*<.001 compared with Checkme at heart level.

[c]*P*=.01 compared with Checkme at heart level in the supine position.

Table 3. Differences in systolic blood pressure readings between the reference monitor and Checkme in various postures and the proportion of volunteers with differences >5, >10, and >15 mmHg between the reference monitor and Checkme.

Reading differences	Supine at heart level (n=37)	Upright at heart level (n=35)
Difference between the devices (mm Hg)		
Mean (SD)	–1.8 (8.5)	2.6 (12.1)[a]
Min; max of range	–19.3; 18.2	–35.5; 20.3
Degree of difference		
>5 mmHg, n (%)	17 (46)	23 (66)
>10 mmHg, n (%)	9 (24)	15 (43)
>15 mmHg, n (%)	5 (14)	6 (17)

[a]*P*=.02 compared with measurements in the supine position.

Figure 5. Bland-Altman plot of the difference in systolic blood pressure readings between the reference monitor (RM) and the Checkme Health Monitor (at heart level) in the supine position.

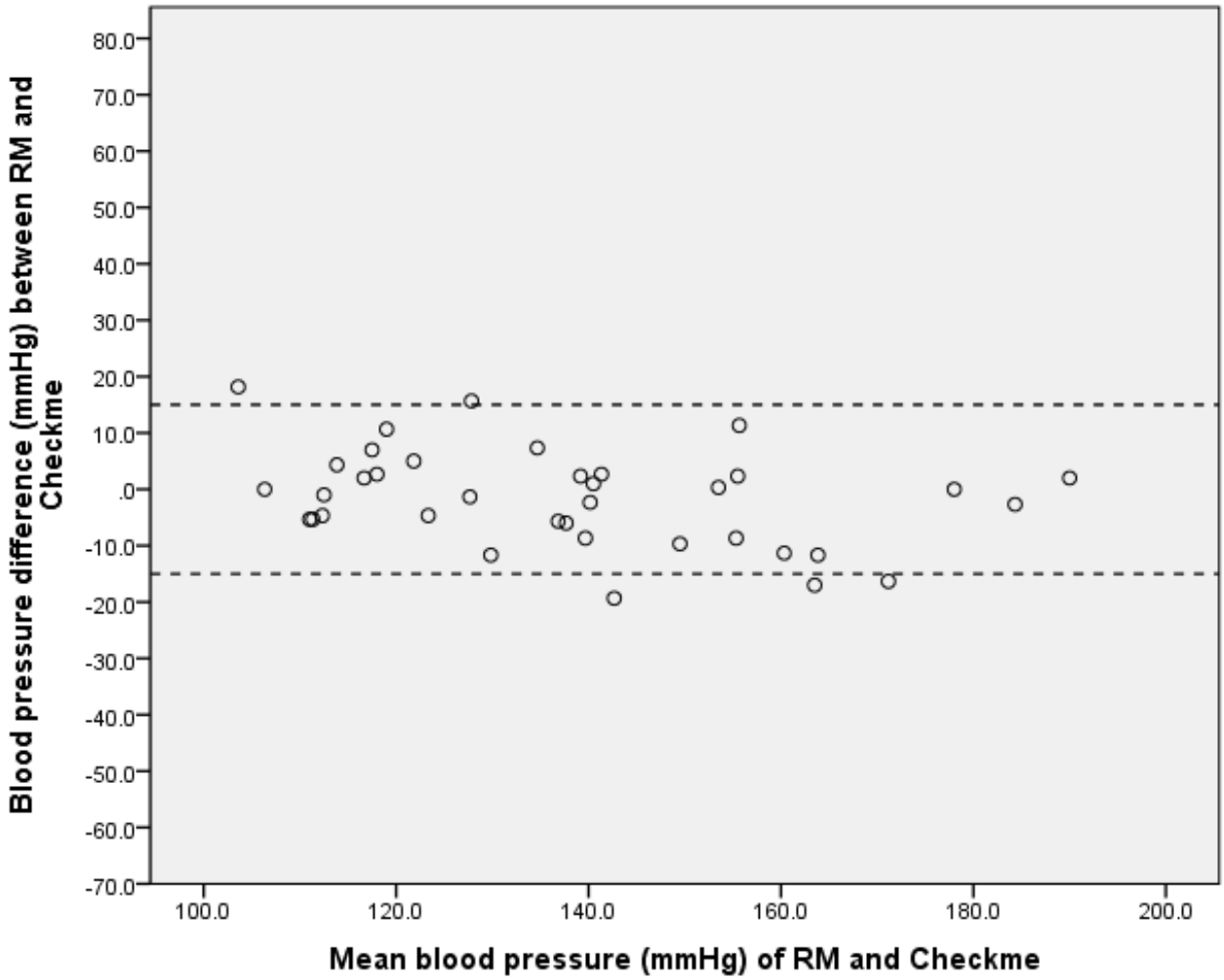
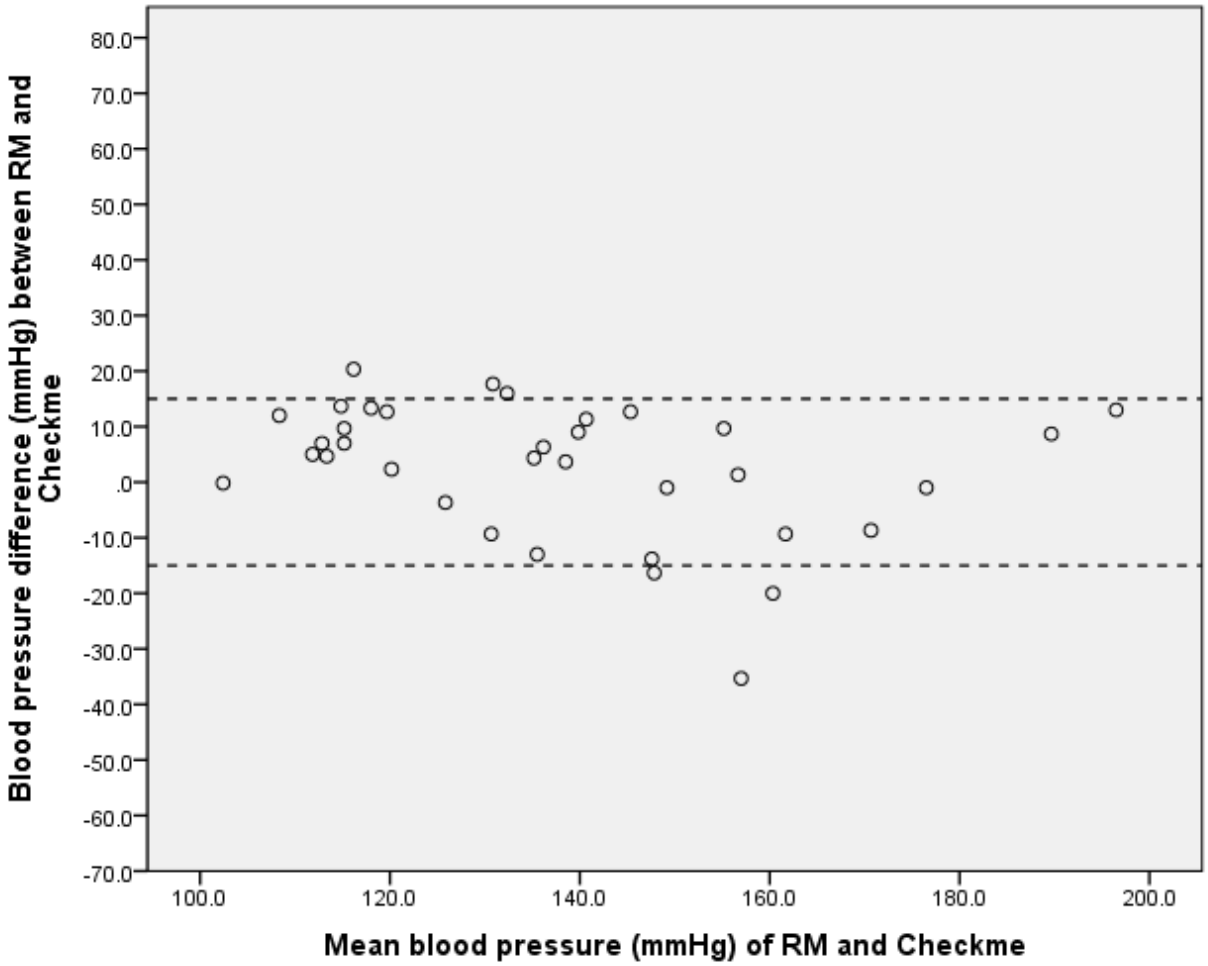


Figure 6. Bland-Altman plot of the difference in systolic blood pressure readings between the reference monitor (RM) and the Checkme Health Monitor (at heart level) in the upright position.



Influence of Posture on the Device Readings (Secondary Aim)

Table 2 summarizes the results of the SBP measurements obtained with RM and Checkme in the various positions. In the supine position, SBP measured with Checkme above heart level was significantly lower than SBP measured supine at heart level. SBP obtained with Checkme in the upright position was significantly lower than in the supine position, in which the device is just above heart level. Table 3 summarizes differences in SBP readings between RM and Checkme in both the supine and upright positions. The SBP measurement with Checkme in the upright position was significantly lower than the SBP measurement with Checkme in the supine position.

Discussion

The results of this comparative study show that the first version of the Checkme device yields BP results that are to a large extend comparable with BPs obtained by a validated oscillometric BP monitor. We observed this for a predefined wide range of BP levels under well-controlled circumstances. Furthermore, BP results correlated with the position of Checkme relative to the heart level. Compared with a reference BP, Checkme recorded a higher BP below heart level and a lower BP above heart level.

Due to the lack of a uniform international protocol that includes provisions to assess intrapatient BP changes relative to the calibrated level, it was not possible to conduct a formal device validation study. As the Checkme requires patient-specific calibration by a secondary measurement device before it can measure absolute BP, we consider such a protocol to be necessary.

The strength of this study is that it met all ESH-IP requirements for test environment, observer qualification, volunteer recruitment, and BP measurements. Measurements were conducted in a quiet, temperature-controlled room and the manufacturer's guidelines on use of the test device were followed. Furthermore, we used a validated RM device and randomized the order of measurements with Checkme and RM to eliminate the influence of changes in BP over time on the study results.

Checkme is one of the first cuffless devices to be launched, indicating that cuffless BP measurement is in its infancy. Notably, Checkme has outgrown its developmental phase. As the technique of cuffless devices is continuously being improved, future generations of Checkme may be even more suitable for measuring BP in the clinic.

One disadvantage of Checkme is the inability to measure DBP, because DBP can be used to calculate pulse pressure and adds

to the overall cardiovascular risk profile. Based on the underlying method of measuring, a subsequent version of Checkme may expected to have this ability. Another issue with the Checkme version used in this study was the inability to calibrate the device in a substantial number of volunteers. Repeated attempts to calibrate Checkme after warming volunteers' hands and further instructing them to hold still or change their position were not effective in some of them and thus further BP measurements were not possible. According to the manufacturer, a new software release has resolved this problem.

Ideally, Checkme is calibrated by taking simultaneous BP measurement with the RM. In this study, we calibrated Checkme after baseline measurements with the RM. However, as the time interval between taking the 2 measurements was a maximum 2 minutes (depending on the number of attempts during calibration), we can assumed that BP had not significantly altered. Calibration parameter stability over longer periods of time has yet to be established in further research. After the completion of this study, Viatom updated the Checkme software to reduce calibration failures and has provided additional instructions for positioning Checkme against a lower limb during the calibration measurements. Therefore, the process of calibration can be expected to be more successful in future studies.

Checkme's BP measuring algorithm has not been made public, probably for commercial reasons. Most cuffless devices measure BP indirectly by determining pulse transit time, the time interval required for a pressure wave in the arterial tree to travel between 2 sites (ie, a proximal and a distal point). Pulse transit time is closely related to BP via arterial compliance. For example, if arterial BP increases, arterial wall tension will increase. Subsequently, arterial compliance and pulse transit time will decrease [9]. Most cuffless devices calculate pulse transit time by using the electrocardiogram as the proximal timing reference and the arterial waveform in an extremity as the distal reference [10]. Recent research has shown a significant relationship between BP measured with pulse transit time and BP measured with conventional devices based on cuff occlusion [11-13].

Differences in BP depending on posture and position of the device suggest that cuffless BP measurement by Checkme, and probably in general, is influenced by the position of the device relative to heart level. This may suggest an inherent error in Checkme's algorithm when BP is measured in a position other than that indicated by the manufacturer. Therefore, it is important that future users of Checkme conduct all measurements in the position stipulated in the user manual. Furthermore, we observed 1 outlier (with SBP difference between RM and Checkme >40 mmHg), which we could explain.

If Checkme will be able to fulfill formal international validation protocol requirements, which include provisions to assess the monitor's accuracy in tracking intrapatient BP changes relative to the calibrated level, after a patient-specific calibration or between calibrations, we expect increased use of this device. Especially promising is such devices' ability to measure BP faster and more conveniently than conventional BP monitoring devices based on cuff occlusion. This implies not only that BP can be measured more efficiently in the clinic, but also that patients can easily self-monitor their BP at home. Because self-measurement of BP has been shown to have a positive effect on reducing BP [3], this easy-to-use BP device will probably find a place in the management of hypertension. The low costs of cuffless devices relative to cuff occlusion devices will also contribute to their implementation in and outside the clinic.

We believe the market of wearable BP sensors will develop in the areas of self-measurement and remote monitoring. In this context, device validation may be accelerated if development of techniques, calculation, and feedback on the basis of clinical data would take place in an open source environment.

Conclusion

Checkme SBP correlated well with reference SBP, in particular in the supine position. Although we did not perform a formal validation study at this preliminary stage, these preliminary results are most promising and warrant further research on cuffless BP measurement in the hospital, the clinic, and at home.

Acknowledgments

We would like to acknowledge Viatom Technology for their entrepreneurial effort in this whole new area. We encourage them to continue their endeavor to improve their sensor devices. We thank Ms MJE Plokker for reviewing the manuscript.

Conflicts of Interest

None declared.

References

1. Perloff D, Grim C, Flack J, Frohlich ED, Hill M, McDonald M, et al. Human blood pressure determination by sphygmomanometry. Circulation 1993 Nov;88(5 Pt 1):2460-2470. [Medline: 8222141]
2. Alpert BS, Quinn D, Gallick D. Oscillometric blood pressure: a review for clinicians. J Am Soc Hypertens 2014 Dec;8(12):930-938. [doi: 10.1016/j.jash.2014.08.014] [Medline: 25492837]
3. 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 6;159(3):185-194. [doi: 10.7326/0003-4819-159-3-201308060-00008] [Medline: 23922064]

4. Omboni S, Gazzola T, Carabelli G, Parati G. Clinical usefulness and cost effectiveness of home blood pressure telemonitoring: meta-analysis of randomized controlled studies. J Hypertens 2013 Mar;31(3):455-67; discussion 467. [doi: 10.1097/HJH.0b013e32835ca8dd] [Medline: 23299557]
5. O'Brien E, Atkins N, Stergiou G, Karpettas N, Parati G, Asmar R, Working Group on Blood Pressure Monitoring of the European Society of Hypertension. European Society of Hypertension International Protocol revision 2010 for the validation of blood pressure measuring devices in adults. Blood Press Monit 2010 Feb;15(1):23-38. [doi: 10.1097/MBP.0b013e3283360e98] [Medline: 20110786]
6. Association for the Advancement of Medical Instrumentation, American National Standards Institute. American National Standard: Non-invasive Sphygmomanometers—Part 1: Requirements and Test Methods for Nonautomated Measurement Type (ANSI/AAMI/ISO 81060-1:2007/(R)2013). Arlington, VA: AAMI; 2008. URL: http://my.aami.org/store/sitecontent/previewfiles/8106001_1306_preview.pdf [accessed 2016-03-28] [WebCite Cache ID 6gLmUcum7]
7. Association for the Advancement of Medical Instrumentation, American National Standards Institute. American National Standard: Non-invasive Sphygmomanometers—Part 2: Clinical Investigation of Automated Measurement Type (ANSI/AAMI/ISO 81060-2:2013). Arlington, VA: AAMI; 2013. URL: http://my.aami.org/aamiresources/previewfiles/8106002_1306_preview.pdf [accessed 2016-03-28] [WebCite Cache ID 6gLmi7fTz]
8. Ng K. Clinical validation protocols for noninvasive blood pressure monitors and their recognition by regulatory authorities and professional organizations: rationale and considerations for a single unified protocol or standard. Blood Press Monit 2013 Oct;18(5):282-289. [doi: 10.1097/MBP.0b013e3283624b3b] [Medline: 23709002]
9. Nichols WW, O'Rourke MF, Vlachopoulos C. McDonald's Blood Flow in Arteries, 6th edition: Theoretical, Experimental and Clinical Principles. Boca Raton, FL: CRC Press; 2011.
10. Kim C, Carek AM, Mukkamala R, Inan OT, Hahn J. Ballistocardiogram as proximal timing reference for pulse transit time measurement: potential for cuffless blood pressure monitoring. IEEE Trans Biomed Eng 2015 Nov;62(11):2657-2664 [FREE Full text] [doi: 10.1109/TBME.2015.2440291] [Medline: 26054058]
11. Masè M, Mattei W, Cucino R, Faes L, Nollo G. Feasibility of cuff-free measurement of systolic and diastolic arterial blood pressure. J Electrocardiol 2011;44(2):201-207. [doi: 10.1016/j.jelectrocard.2010.11.019] [Medline: 21353067]
12. Gesche H, Grosskurth D, Küchler G, Patzak A. Continuous blood pressure measurement by using the pulse transit time: comparison to a cuff-based method. Eur J Appl Physiol 2012 Jan;112(1):309-315. [doi: 10.1007/s00421-011-1983-3] [Medline: 21556814]
13. Ma HT. A blood pressure monitoring method for stroke management. Biomed Res Int 2014;2014:571623 [FREE Full text] [doi: 10.1155/2014/571623] [Medline: 25197651]

Abbreviations

BP: blood pressure
DBP: diastolic blood pressure
ESH-IP: European Society of Hypertension International Protocol
RM: reference blood pressure monitor
SBP: systolic blood pressure

Edited by G Eysenbach; submitted 07.12.15; peer-reviewed by KY Chin, N Farič; comments to author 14.01.16; revised version received 05.02.16; accepted 20.02.16; published 05.05.16.

Please cite as:
Schoot TS, Weenk M, van de Belt TH, Engelen LJLPG, van Goor H, Bredie SJH
A New Cuffless Device for Measuring Blood Pressure: A Real-Life Validation Study
J Med Internet Res 2016;18(5):e85
URL: http://www.jmir.org/2016/5/e85/
doi:10.2196/jmir.5414
PMID:27150527

Original Paper

Patient-Reported Safety Events in Chronic Kidney Disease Recorded With an Interactive Voice-Inquiry Dial-Response System: Monthly Report Analysis

Jeffrey C Fink[1], MD, MS; Rebecca M Doerfler[1], MPH; Marni R Yoffe[1], MA; Clarissa J Diamantidis[2], MD, MHS; Jacob B Blumenthal[1], MD; Tariq Siddiqui[1], MIS; James F Gardner[3], ScM; Soren Snitker[1], MD, PhD; Min Zhan[4], PhD

[1]University of Maryland School of Medicine, Department of Medicine, Baltimore, MD, United States

[2]Duke University School of Medicine, Department of Medicine, Durham, NC, United States

[3]University of Maryland School of Pharmacy, Pharmaceutical Research Computing, Baltimore, MD, United States

[4]University of Maryland, Department of Epidemiology and Public Health, Baltimore, MD, United States

Corresponding Author:
Jeffrey C Fink, MD, MS
University of Maryland School of Medicine
Department of Medicine
22 S Greene St
Baltimore, MD, 21201
United States
Phone: 1 4107066563
Fax: 1 4107066562
Email: jfink@medicine.umaryland.edu

Abstract

Background: Monitoring patient-reported outcomes (PROs) may improve safety of chronic kidney disease (CKD) patients.

Objective: Evaluate the performance of an interactive voice-inquiry dial-response system (IVRDS) in detecting CKD-pertinent adverse safety events outside of the clinical environment and compare the incidence of events using the IVDRS to that detected by paper diary.

Methods: This was a 6-month study of Stage III-V CKD patients in the Safe Kidney Care (SKC) study. Participants crossed over from a paper diary to the IVDRS for recording patient-reported safety events defined as symptoms or events attributable to medications or care. The IVDRS was adapted from the SKC paper diary to record event frequency and remediation. Participants were auto-called weekly and permitted to self-initiate calls. Monthly reports were reviewed by two physician adjudicators for their clinical significance.

Results: 52 participants were followed over a total of 1384 weeks. 28 out of 52 participants (54%) reported events using the IVDRS versus 8 out of 52 (15%) with the paper diary; hypoglycemia was the most common event for both methods. All IVDRS menu options were selected at least once except for confusion and rash. Events were reported on 121 calls, with 8 calls reporting event remediation by ambulance or emergency room (ER) visit. The event rate with the IVDRS and paper diary, with and without hypoglycemia, was 26.7 versus 4.7 and 18.3 versus 0.8 per 100 person weeks, respectively (P=.002 and P<.001). The frequent users (ie, >10 events) largely differed by method, and event rates excluding the most frequent user of each were 16.9 versus 2.5 per 100 person weeks, respectively (P<.001). Adjudicators found approximately half the 80 reports clinically significant, with about a quarter judged as actionable. Hypoglycemia was often associated with additional reports of fatigue and falling. Participants expressed favorable satisfaction with the IVDRS.

Conclusions: Use of the IVDRS among CKD patients reveals a high frequency of clinically significant safety events and has the potential to be used as an important supplement to clinical care for improving patient safety.

(J Med Internet Res 2016;18(5):e125) doi:10.2196/jmir.5203

KEYWORDS

patient-reported outcomes; CKD; interactive voice-response system; patient safety

Introduction

Recording patient-reported outcomes (PROs) is important for effective management of chronic disease management. The US Food and Drug Administration (FDA) defined a PRO as a report of health status coming directly from a patient "without amendment or interpretation by a clinician" [1]. The nephrology community recognizes the need to monitor PROs in chronic kidney disease (CKD) care [2,3]. Areas of emphasis to date include assessing dialysis patients' physical and mental impairment [4], preferences and experiences with renal replacement therapy [5,6], and the symptom burden of CKD-related anemia [7]. General tools for documenting PROs, such as the 36- and 12-item Short Form Health Survey (SF-36 and SF-12,respectively), are used in kidney disease; however, kidney disease PRO assessments are more commonly applied in the research domain than in the clinic [8]. Tools to record clinically significant and actionable CKD PROs are needed [2].

Monitoring and interpreting PROs may give insight into means of reducing patient safety events, here defined as harm from medical therapy [9-11]. Presenting underdetected PROs to providers, especially medication-related adverse events, can offer new opportunities to improve the safety of delivered care. Patient diaries have been employed to record PROs related to medication tolerance in clinical trials as well as with disease management, and both paper and electronic mediums have been used [12,13]. CKD patients have multiple comorbidities, take numerous medications, often report safety-related outcomes and experiences [14], and may benefit from monitoring such PROs.

In this study, we examined a subcohort of predialysis CKD patients in the Safe Kidney Care (SKC) study (NCT01407367). We used an interactive voice-inquiry dial-response system (IVDRS) to gauge patient-reported experiences attributed to medications and considered adverse safety events. We demonstrate the participants' usage of the IVDRS relative to the SKC paper diary, report the serial trends in engagement, tabulate the types of events reported, evaluate the clinical importance of the reported events, and survey participant satisfaction with the reporting system.

Methods

Study Overview

The IVDRS study is an ancillary to the ongoing SKC cohort study with the latter commencing in 2011. The protocol and informed consent were approved by the University of Maryland Institutional Review Board. The SKC study tracked CKD patients through annual in-center visits to determine the frequency of an array of adverse safety events. The study included two subgroups: Phase 1 participants were provided with a medical alert accessory to augment kidney disease awareness and were given access to the SKC website providing best practices in safe CKD care; Phase 2 SKC participants received no accessory or website access, but were followed on the same schedule as Phase 1 participants for detection of adverse safety events. SKC monitoring included provision of a paper diary to document adverse events that the participant attributed to a medication, medical instructions, or medical care. The structured check-off entries in the SKC diary were predetermined after an online nephrologist survey and an expert panel reviewed, adjudicated, and categorized prominent adverse safety events. The diary also permitted text entries not included in the menu of structured events. Participants were instructed to mail in diary pages—using study-issued postage-paid envelopes—with documented entries and, as per the protocol, were reminded every 3 months to use their diary. Text entries were reviewed semiannually to determine if they could be reclassified into a structured entry or if they warranted a unique adverse safety event category.

The IVDRS platform was a Linux-based operating system provided and administered by CircleLink Health (Stamford, CT). The phone-based interactive modality uses the telephone dial-pad for participant responses to automated voice queries. The IVDRS query menu is derived from the SKC paper diary using a similarly structured list of adverse safety events, but with each adverse safety event linked to a series of more-detailed questions, including event frequency and determination of the action taken for each event (eg, "Did you feel dizzy in the last 7 days? Press '1' for yes or '2' for no." "Enter the number of times you felt dizzy. Was the dizziness caused by a medicine? Press '1' for yes or '2' for no." "How did you treat your dizziness? Press '1' if you called an ambulance or went to the emergency room. '2' if you called your doctor. '3' if you self-treated. '4' if you did nothing. '5' if you did more than one of these."). The IVDRS query for hypoglycemia included a threshold of less than 70 mg/dL versus 60 mg/dL by paper diary. The more stringent threshold with the latter was based on the expectation that participants would be less inclined to mail in entries for less severe hypoglycemia. Given the number of potential adverse safety events, the first menu of dial options included what was expected to be the most frequent events along with an option to proceed to a second menu for the remaining adverse safety events to select if they occurred. An option was also included to dial in if no events occurred. The IVDRS protocol was programmed to call study participants weekly and, once calls were answered, the IVDRS queried participants about events during that day and the preceding 6 days. If a given weekly call was unanswered, the system would call again 15 minutes later, but if there was no answer on the second call, then the IVDRS ceased calling until the next week. The IVDRS also had the capability for participants to initiate calls for events they deemed necessary to document prior to the scheduled weekly call with an associated time stamp. If the participant elected to use the call-in option within 24 hours of the scheduled weekly call, the latter was not initiated.

Study Participants and Baseline Assessment

The IVDRS study enrolled 52 consecutive consenting SKC participants between June 2, 2014, and December 18, 2014, at any postbaseline annual study visit, without regard to the frequency of paper diary use. Participants without a stable telephone number—cellular or ground line—or who were expected to reach dialysis or die over the next 6 months were not enrolled in the IVDRS study. Information collected from the core SKC study visit when the participant was recruited included the following: in-center vital signs, serum 6-8-hour

fasting glucose, potassium, creatinine for estimation of glomerular filtration rate (GFR) and venous hemoglobin, and the annual medications reported. Also, paper diary pages from the prior 6 months were reviewed and events were documented.

Each consenting participant was asked for a preferred weekly IVDRS call time over the 6 months of the protocol. Each participant was instructed on the use of the IVDRS system with a mock IVDRS interaction and was told that events to be submitted were those they believed to be attributable to a medicine or their medical care, similar to the SKC paper diary protocol. At the completion of the protocol, each participant returned to the study center for an update of reported medications, medical event update, and satisfaction survey.

IVDRS data was received by the vendor and transmitted to the study team via a Web-based portal with daily coordinator review for obvious data entry errors; urgent adverse safety events (eg, ≥ 1 reported fall) were flagged and promptly reviewed. A monthly report was prepared for each participant with all events detected via weekly calls or self-initiated calls in a 4-week window. Monthly reports along with the SKC visit vital sign readings, laboratory values, medical comorbidities, and medications were presented to two independent physician reviewers (JBB, CJD) for determination of whether each report was (1) of no interest, (2) of interest, but with no action recommended, or (3) of interest with action recommended. A separate physician reviewer (SS) examined each adverse safety event to determine if participants with adverse events had reported taking a medication that could plausibly cause such an experience based on attributable medication categories previously defined [14].

Analysis

The analysis was designed to be descriptive with demographic characteristics of the IVDRS participants reported as means (SD) and with n (%) for categorical variables. The IVDRS patient-reported adverse safety events were compared to the SKC paper diary patient reports as a baseline reference and using the generalized estimating equation (GEE) with Poisson distribution and a log link function. IVDRS and SKC paper diary event rates were reported per 100 participant weeks; IVDRS adverse events were also presented by months of participation and by adverse safety event category. The final visit was 26 ± 2 weeks in length to accommodate participants' schedules. Association (ie, market basket) analysis was also employed to identify the likelihood of coexistent adverse safety event types by participant using the IVDRS and treating the entire study period as a single observation period. Association analysis measures included confidence, which indicates the likelihood of a consequent event given the occurrence of an antecedent event, and the lift ratio, which is the confidence over the prevalence of the consequent event independent of the antecedent. A lift ratio of 1 indicated that the co-occurrence of events was random [14-16]. Analyses were conducted in SAS version 9.3 (SAS Institute, Inc; Cary, NC).

Results

Characteristics of Study Participants

The characteristics of the 52 participants are shown in Table 1. Almost 60% of participants were greater than 65 years of age and a majority were male and African American. Approximately two-thirds of participants had Stage III-B through Stage V CKD at the time of enrollment, most had diabetes, and about half had previously used the Internet. The averages of study participants' clinical measures were in acceptable ranges; however, there was substantial polypharmacy. All study participants completed the study protocol, except for one who died after completing 22 weeks of the 6-month protocol.

Table 1. Characteristics of interactive voice-inquiry dial-response system (IVDRS) study participants.

Participant characteristics	Participants (n=52) or values, n (%), mean (SD), or median (Interquartile range)
Time in Safe Kidney Care study, n (%)	
1 year	21 (40)
2 years	14 (27)
3 years	17 (33)
Age in years, n (%)	
≤ 60	10 (19)
61-65	11 (21)
66-70	18 (35)
≥71	13 (25)
Gender, n (%)	
Male	40 (77)
Female	12 (23)
African American, n (%)	
Yes	39 (75)
No	13 (25)
Systolic blood pressure (mm Hg), mean (SD)	
Sitting	134 (20.6)
Standing, (n=50), 2 participants could not stand	134 (22.4)
Diastolic blood pressure (mm Hg), mean (SD)	
Sitting	70 (12.6)
Standing (n=50), 2 participants could not stand	72 (15.5)
Glomerular filtration rate (ml/min/1.73m ^{2}), n (%)	
≤ 45	34 (65)
> 45	18 (35)
Serum potassium (meq/L), mean (SD)	4.4 (0.6)
Serum glucose (mg/dL), mean (SD)	136.3 (57.4)
Venous hemoglobin (g/dL), mean (SD)	11.9 (1.6)
Cardiovascular disease, n (%)	
Yes	23 (44)
No	29 (56)
Cancer, n (%)	
Yes	17 (33)
No	35 (67)
Diabetes, n (%)	
Yes	47 (90)
No	5 (10)
Used the Internet in the past year to look for health information, n (%)	
Yes	23 (44)
No	29 (56)
Number of medications, median (IQR)	16 (10.0)

Use of the Interactive Voice-Inquiry Dial-Response System

Table 2 shows study participation over 1384 weeks distributed by reporting any or no safety events and further classified by expected and detected weeks of participation. Table 2 also tabulates call response by frequency of events per call and participant remediation of events reported on each call. A total of 24 out of 52 (46%) study participants reported no events over 650 weeks of participation, including 619 weeks when a call was delivered or self-initiated. A total of 28 out of 52 (54%) study participants reported at least one safety event over 734 total weeks of participation, of which 731 had calls delivered or self-initiated. Events were reported in 113 weeks with a total of 121 calls either delivered or self-initiated with one or more reported events. Most calls included one event and self-treatment, but a notable minority (8/121, 6.6%) involved an ambulance or emergency room (ER) visit.

Figure 1 shows the trend in monthly study engagement along with event reporting. The dashed line shows the trend in the number of active participant weeks per month—an active week is defined as one or more dialed entry, including "no events to report"—and the solid line represents total number of participant weeks with safety events per month. While the frequency of event reporting declined over the study, the IVDRS engagement measured by active participant weeks remained relatively constant.

Only 10 out of 1227 (0.81%) calls with participant entries were found to be erroneous after coordinator-initiated call inspection or by system alert for extreme values. In such instances, staff contacted participants and database corrections were made after clarification. A total of 31 out of 1227 (2.53%) initial entries had errors that were self-corrected during the call when the participant was prompted to confirm his/her selection.

Table 2. Distribution of interactive voice-inquiry dial-response system activity by participants, weeks of participation, and events.

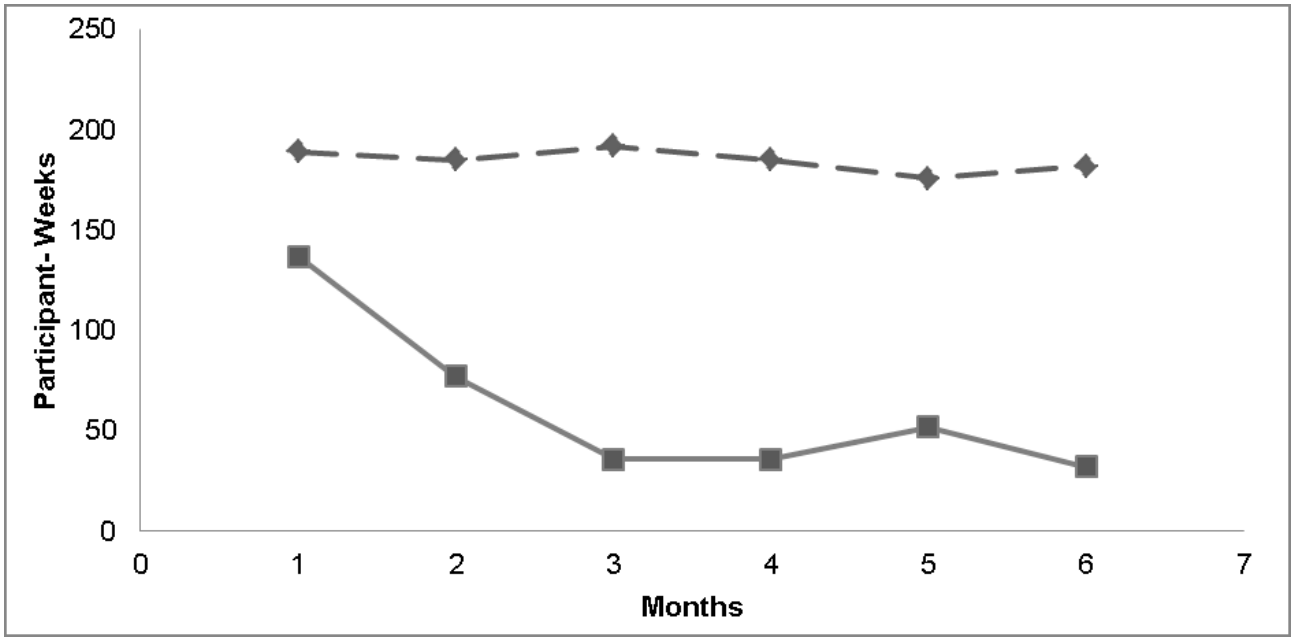
Participant type	Number of participants (n=52), n (%)	Weeks of participation (n=1384), n (%)	Weeks with delivered or self-initiated calls (n=1350), n (%)	Calls with events reported (n=121)[a], n: number of events or remediation	
				Calls by number of events	Calls by associated remediation
With no events	24 (46)				
No call		31 (2.24)			
Call delivered or self-initiated		619 (44.73)			
Not picked up			80 (5.93)		
No events reported			539 (39.93)		
With events	28 (54)				
No call		3 (0.22)			
Call delivered or self-initiated		731 (52.82)			
Not picked up			138 (10.22)		
No events reported			480 (35.56)		
Events reported			113 (8.37)		
Event details				115: 1 event 5: 2 events 1: 3 events	8: Ambulance/ER[b] visit 3: called MD[c] 90: self-treated 14: nothing done 4: multiple actions 1: hung up

[a]Participants may self-initiate a call more than once a week.

[b]ER: emergency room.

[c]MD: medical doctor.

Figure 1. Participant weeks with study participation and adverse event reporting by month. The dashed line shows the trend in the number of active participant weeks per month—an active week is defined as one or more dialed entry, including "no events to report"—and the solid line represents total number of participant weeks with safety events per month.



Patient-Reported Safety Events Using the Interactive Voice-Inquiry Dial-Response System Relative to the Paper Diary

Table 3 shows that the number of reported events was significantly greater with the IVDRS versus the paper diary. In the 6 months preceding IVDRS enrollment, participants submitted 85 paper diary entries with 95 patient-reported events. Of those 95 events, 39 (41%) patient-reported symptoms were text entries, with 3 out of 39 (8%) classified into existing categories and 5 out of 39 (13%) reports of feeling sick or having bruises, which were classified as other adverse safety events. A total of 31 out of 95 (33%) text entries were for pain and were considered as not meeting the criteria of patient-reported safety events (ie, not related to medical care). The exclusion of these entries reduced the total number of patient-reported safety events to 64.

Table 3. Comparison of adverse safety events detected by the Safe Kidney Care paper diary versus the interactive voice-inquiry dial-response system (IVDRS).

Participant categories	Paper diary, n or n (participant [P] identifier)	IVDRS[a], n or n (participant [P] identifier)
Total weeks of participation[b]	1352	1384
Participants with events	8	28
Participants with no events	44	24
Total safety events	64	370
Maximum safety events per participant	32	101
Total safety events excluding highest reporter	32	225
Total events excluding hypoglycemia	11	253
Events per 100 participant weeks	4.7	26.7[c]
Events per 100 participant weeks excluding highest user	2.5	16.9[d]
Events per 100 participant weeks excluding hypoglycemia	0.8	18.3[d]
Top reporters and number of events, n (participant [P] identifier)		
1	32 [P5]	101 [P33]
2	12 [P3]	45 [P27]
3	10 [P29]	44 [P5]
4	5 [P50]	32 [P20]
5	2 [P4]	28 [P41]
6	1 [P1, P9, P10]	26 [P50]
Top safety events reported		
Hypoglycemia	53	117
Leg swelling	1	80
Dizziness	1	77
Falling	0	12
Bleeding	0	1
Other	9	83

[a]IVDRS: interactive voice-inquiry and dial-response system.

[b]One participant did not complete the study but contributed 22 weeks.

[c]P=.002.

[d]P<.001.

With both modalities, hypoglycemia was the most common safety event, but the number of reported hypoglycemic events was greater with the IVDRS than with the paper diary. When excluding the most frequent reporter of adverse safety events with each modality, the frequency of reporting with IVDRS remained significantly higher. Similarly, when excluding hypoglycemia, the IVDRS modality had a significantly higher rate of reported events. Table 3 also displays the top 6 reporters by modality and demonstrates that 4 frequent users in each modality were not among the top-ranked users with the other modality.

Table 4 reveals the distribution of all 370 events reported via the IVDRS by study month. Hypoglycemia represented 31.6% of events across the study period, followed by leg or ankle swelling, dizziness, and fatigue. Of note, only one participant reported events that were not conditional on a potentially attributable medication. Omitting those events not meeting this conditional criterion for a safety event reduced the total number of events from 370 to 356.

Table 4. Distribution of patient-reported adverse safety events using the interactive voice-inquiry dial-response system (IVDRS) over the study period.

Symptoms	Patient-reported adverse safety events, n (%)						
	Months						
	1 (n=137)	2 (n=77)	3 (n=36)	4 (n=36)	5 (n=52)	≥6 (n=32)	Total (n=370)
Total symptoms (n=370)	137	77	36	36	52	32	370
Low blood sugar	27 (19.7)	26 (34)	15 (42)	24 (67)	10 (19)	15 (47)	117 (31.6)
Leg or ankle swelling	27 (19.7)	29 (38)	11 (31)	0 (0)	12 (23)	1 (3)	80 (21.6)
Face swelling	0 (0)	7 (9)	0 (0)	0 (0)	0 (0)	0 (0)	7 (1.9)
Dizziness	28 (20.4)	5 (6)	0 (0)	0 (0)	28 (54)	16 (48)	77 (20.8)
Fall	1 (0.7)	0 (0)	3 (8)	7 (19)	1 (2)	0 (0)	12 (3.2)
Stomach problems	7 (5.1)	2 (3)	0 (0)	0 (0)	0 (0)	0 (0)	9 (2.4)
Bleeding	0 (0)	0 (0)	0 (0)	0 (0)	1 (2)	0 (0)	1 (0.3)
High potassium	1 (0.7)	1 (1)	0 (0)	0 (0)	0 (0)	0 (0)	2 (0.5)
Skin rash	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Confusion	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Fatigue	46 (33.6)	7 (9)	7 (19)	0 (0)	0 (0)	0 (0)	60 (16.2)
Low blood pressure	0 (0)	0 (0)	0 (0)	5 (14)	0 (0)	0 (0)	5 (1.4)

With 23 of 52 participants (44%) reporting more than one event during the study period, we found the strongest association of hypoglycemia with fatigue and falling. Participants who reported fatigue had 66.6% confidence of reporting an episode of hypoglycemia at some time during the study period, with a lift ratio of 2.31. Likewise, participants who reported a fall had 50% confidence of reporting a hypoglycemic episode resulting in a lift ratio of 1.73.

Clinical Significance of Patient-Reported Safety Events

A significant proportion of patient-reported safety events were judged to be clinically important by the pair of physician adjudicators reviewing participants' cumulative monthly reports of all IVDRS-reported events, in conjunction with SKC-measured laboratory values—estimated GFR (eGFR), potassium, hemoglobin, and fasting glucose—reported comorbidities, and medications collected on study entry. Of the 80 reports reviewed, the two adjudicators individually found 20.0% and 11.3% of them of no clinical interest, 53.8% and 46.3% of them of clinical interest but no action would be taken, and 26.3% and 42.5% of them warranting of clinical action.

Participant Feedback and Satisfaction With the Interactive Voice-Inquiry Dial-Response System

At completion of the study, participants reported a universally high degree of satisfaction with the IVDRS. Table 5 shows that more than 90% of the participants agreed or strongly agreed that the IVDRS was easy to use and was used confidently and weekly. Most participants liked using the IVDRS and would recommend its use to other CKD patients in the future.

Table 5. Participant satisfaction with interactive voice-inquiry and dial-response system (IVDRS) use.

Satisfaction survey items	Participant response (n=51[a]), n (%)						
	1 (strongly agree)	2	3	4 (neutral)	5	6	7 (strongly disagree)
It is easy to use the eDiary to record safety events that happen to me	22 (43)	28 (55)	0 (0)	1 (2)	0 (0)	0 (0)	0 (0)
I would recommend the eDiary service to other people with kidney problems	22 (43)	27 (53)	0 (0)	1 (2)	0 (0)	1 (2)	0 (0)
I like using the eDiary to record safety events that happen to me	20 (39)	27 (53)	1 (2)	1 (2)	0 (0)	2 (4)	0 (0)
I think that I would like to use this service often to report medical safety events to my doctors	20 (39)	27 (53)	0 (0)	1 (2)	0 (0)	3 (6)	0 (0)
I like when the system called me on a weekly basis to ask me to record medical safety events	18 (35)	30 (59)	0 (0)	1 (2)	1 (2)	1 (2)	0 (0)
I felt very confident using the eDiary	17 (33)	31 (61)	1 (2)	0 (0)	1 (2)	1 (2)	0 (0)
I felt that the eDiary system was hard to use	0 (0)	0 (0)	1 (2)	0 (0)	0 (0)	37 (73)	13 (26)

[a]One participant died before completion of the protocol.

Discussion

Principal Findings

PROs can provide important information on CKD-related adverse safety events outside the medical care setting. Such experiences may be overlooked or underreported because they are not recalled by patients or solicited by providers; moreover, patients may have difficulty in drawing associations between such events and their treatments. This study demonstrates the utility of an IVDRS in monitoring a dimension of the patient experience that can be related to safe care. The application of this elementary form of remote data capture of CKD patient-reported safety incidents reveals a high rate of events in the context of potentially attributable medication usage. The number of events captured by the IVDRS is high relative to a paper diary used by the same participants. A notable number of the reported events were found to be significant as they led to urgent responses, while conversely, a high frequency of these PROs were self-treated and perhaps underappreciated by the reporters for their clinical significance. Independent review of these events revealed a substantial proportion of them to be important and actionable given the clinical context in which they occurred. Participants were favorable in their review of the system and provided encouraging evidence that this common household technology can be used to enhance chronic disease management in CKD [17].

Limitations

The study has inherent limitations to be considered when interpreting the results. The completeness of data collected using the IVDRS to record safety events is limited by participants' motivation in using the system. We attempted to minimize potential reporting bias based on severity or frequency of events by scheduling weekly calls and limiting participant dependence on memory. The relatively short time frame between calls was intended to improve recollection of events across a wider severity range and avoid restricting the selection of only the most significant of events.

While the comparisons between the IVDRS and paper diary means of event recording are paired within participants, the modalities have inherent differences that must be acknowledged when contrasting event rates. The scheduled contact of participants by the IVDRS effectively solicits responses in comparison to the expected self-initiated use of the paper diary—even with 3-month reminders—and most likely enhances the detection of adverse events. The IVDRS menu of delivered questions may compel the participant to consider more thoroughly their safety experiences over the duration of the study. Additionally, the higher threshold set with the IVDRS for incidence of hypoglycemia may also trigger a higher reporting rate relative to the paper diary. However, examining the frequency of events with each modality is still informative in demonstrating the extent to which the IVDRS detects a wide range and high frequency of adverse safety events, even when excluding hypoglycemia.

The IVDRS protocol resulted in an apparent time-dependent decline in patient-reported safety events. While the secular fall-off in event reporting could be considered the result of user fatigue over the duration of the protocol, engagement did not similarly decline over the study period, suggesting that participants did not lose interest in the IVDRS. This finding raises the possibility that there was a learning effect where adverse safety events truly declined as participants gained self-management abilities through the protocol. Finally, the reporting of adverse safety events relies on the participants' judgement of what was attributable to medications and treatments, and raises the possibility of inaccurate reporting. Nevertheless, additional physician review demonstrated the vast majority of events reported could plausibly be related to medications administered at the time of the event. This review reduced the likelihood of overreporting, but did not diminish the possibility that participants may have failed to report events

that could have been attributed to a medication, but they judged otherwise.

Comparison With Prior Work

Diary methods have become increasingly used to record PROs. Use of Internet and paper-based diaries have become widely prevalent and range from monitoring of headaches, epilepsy activity, rheumatoid arthritis severity, and glucose control in diabetes [18-21]. These methods include acquisition of real-time experiences or ecological momentary assessments using end-of-day evaluations or the Day Reconstruction Method [12,13,22]. These methodologies can gauge participant sentiments, quality-of-life estimates, or symptoms as proposed here [12]. Such diaries are subject to recall bias, imprecise measurements, and potential overweighting of negative experiences [4,23].The use of newer hand-held electronic devices to record PROs needs to overcome challenges related to security, cost, and technological limitations of the target population, as are common in CKD and dialysis populations [13].

We are unaware of prior studies using diaries to record patient-related safety experiences in CKD. We elected to use a telephone-based portal, typically using a landline, to engage CKD patients in recording experiences related to symptoms or incidents they viewed as attributable to their medical treatment. This communication modality matches the technological proficiency characteristics of much of the target population [24,25]. We have previously shown that CKD patients have variable success with the use of mobile devices and Web-based applications [24,25]. We also chose a weekly scheduled call in order to balance bias anticipated from a 7-day recall period with the inconvenience of more frequent scheduled calls.

The study also demonstrates the prominence of hypoglycemia in the CKD population where diabetes is common along with the use of insulin and other diabetic treatments. However, the rate of nonhypoglycemic events with the IVDRS, as opposed to with the SKC paper diary, reveals the broader range of PROs detected with this technology. The association of several patient-reported experiences with hypoglycemia corroborates findings from a prior report linking hypoglycemia to other patient-reported safety incidents [14].

Conclusions

We have demonstrated that an IVDRS designed to detect adverse safety events sheds light on a broader dimension of the CKD “phenotype” that is likely to be underreported in the clinical setting. The increased detection of PROs using a low-cost and simple communication technology has the potential to enhance the care of CKD patients, improve their safety, and address the high rate of poor outcomes common in this population.

Acknowledgments

The work included in this paper was supported by the National Institute of Diabetes and Digestive and Kidney Disease (R01-DK084017 and R21-DK096204). The IVDRS was developed and administered by CircleLink (Stamford, CT).

Conflicts of Interest

None declared.

References

1. Guidance for Industry. Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims. Silver Spring, MD: US Department of Health and Human Services, Food and Drug Administration; 2009 Dec. URL: http://webcache.googleusercontent.com/search?q=cache:http://www.fda.gov/downloads/Drugs/.../Guidances/UCM193282.pdf&gws_rd=cr&ei=-74sV9yFIoy8eY-wpcgH [accessed 2016-05-06] [WebCite Cache ID 6hJ1FXccs]
2. Perrone RD, Coons SJ, Cavanaugh K, Finkelstein F, Meyer KB. Patient-reported outcomes in clinical trials of CKD-related therapies: Report of a symposium sponsored by the national kidney foundation and the US Food and Drug Administration. Am J Kidney Dis 2013 Dec;62(6):1046-1057. [doi: 10.1053/j.ajkd.2013.07.004] [Medline: 23988757]
3. Breckenridge K, Bekker HL, Gibbons E, van der Veer SN, Abbott D, Briançon S, et al. How to routinely collect data on patient-reported outcome and experience measures in renal registries in Europe: An expert consensus meeting. Nephrol Dial Transplant 2015 Oct;30(10):1605-1614 [FREE Full text] [doi: 10.1093/ndt/gfv209] [Medline: 25982327]
4. Mittal SK, Ahern L, Flaster E, Maesaka JK, Fishbane S. Self-assessed physical and mental function of haemodialysis patients. Nephrol Dial Transplant 2001 Jul;16(7):1387-1394 [FREE Full text] [Medline: 11427630]
5. Barendse SM, Speight J, Bradley C. The Renal Treatment Satisfaction Questionnaire (RTSQ): A measure of satisfaction with treatment for chronic kidney failure. Am J Kidney Dis 2005 Mar;45(3):572-579. [doi: 10.1053/j.ajkd.2004.11.010] [Medline: 15754280]
6. Rubin HR, Jenckes M, Fink NE, Meyer K, Wu AW, Bass EB, et al. Patient's view of dialysis care: Development of a taxonomy and rating of importance of different aspects of care. CHOICE study. Choices for Healthy Outcomes in Caring for ESRD. Am J Kidney Dis 1997 Dec;30(6):793-801. [Medline: 9398123]
7. Finkelstein FO, Story K, Firanek C, Mendelssohn D, Barre P, Takano T, et al. Health-related quality of life and hemoglobin levels in chronic kidney disease patients. Clin J Am Soc Nephrol 2009 Jan;4(1):33-38 [FREE Full text] [doi: 10.2215/CJN.00630208] [Medline: 18987300]

8. Lacson E, Xu J, Lin S, Dean SG, Lazarus JM, Hakim RM. A comparison of SF-36 and SF-12 composite scores and subsequent hospitalization and mortality risks in long-term dialysis patients. Clin J Am Soc Nephrol 2010 Feb;5(2):252-260 [FREE Full text] [doi: 10.2215/CJN.07231009] [Medline: 20019120]
9. Basch E, Bennett A, Pietanza MC. Use of patient-reported outcomes to improve the predictive accuracy of clinician-reported adverse events. J Natl Cancer Inst 2011 Dec 21;103(24):1808-1810 [FREE Full text] [doi: 10.1093/jnci/djr493] [Medline: 22157639]
10. Basch E. The missing voice of patients in drug-safety reporting. N Engl J Med 2010 Mar 11;362(10):865-869 [FREE Full text] [doi: 10.1056/NEJMp0911494] [Medline: 20220181]
11. 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]
12. Schneider S, Stone AA. Ambulatory and diary methods can facilitate the measurement of patient-reported outcomes. Qual Life Res 2016 Mar;25(3):497-506. [doi: 10.1007/s11136-015-1054-z] [Medline: 26101141]
13. Schick-Makaroff K, Molzahn A. Strategies to use tablet computers for collection of electronic patient-reported outcomes. Health Qual Life Outcomes 2015;13:2 [FREE Full text] [doi: 10.1186/s12955-014-0205-1] [Medline: 25608560]
14. Ginsberg JS, Zhan M, Diamantidis CJ, Woods C, Chen J, Fink JC. Patient-reported and actionable safety events in CKD. J Am Soc Nephrol 2014 Jul;25(7):1564-1573 [FREE Full text] [doi: 10.1681/ASN.2013090921] [Medline: 24556352]
15. Tian H, Brimmer DJ, Lin JS, Tumpey AJ, Reeves WC. Web usage data as a means of evaluating public health messaging and outreach. J Med Internet Res 2009;11(4):e52 [FREE Full text] [doi: 10.2196/jmir.1278] [Medline: 20026451]
16. Wang X. SAS Global Forum. 2010. Investigating the impact of Medicare, Part D on the diabetes medications using Enterprise Miner and survival analysis URL: http://support.sas.com/resources/papers/proceedings10/126-2010.pdf [accessed 2016-04-29] [WebCite Cache ID 6h8LSFfqo]
17. Tabak M, op den Akker H, Hermens H. Motivational cues as real-time feedback for changing daily activity behavior of patients with COPD. Patient Educ Couns 2014 Mar;94(3):372-378. [doi: 10.1016/j.pec.2013.10.014] [Medline: 24332934]
18. Krogh A, Larsson B, Salvesen Ø, Linde M. A comparison between prospective Internet-based and paper diary recordings of headache among adolescents in the general population. Cephalalgia 2016 Apr;36(4):335-345. [doi: 10.1177/0333102415591506] [Medline: 26092285]
19. Hall CB, Lipton RB, Tennen H, Haut SR. Early follow-up data from seizure diaries can be used to predict subsequent seizures in same cohort by borrowing strength across participants. Epilepsy Behav 2009 Mar;14(3):472-475 [FREE Full text] [doi: 10.1016/j.yebeh.2008.12.011] [Medline: 19138755]
20. Heiberg T, Kvien TK, Dale Ø, Mowinckel P, Aanerud GJ, Songe-Møller AB, et al. Daily health status registration (patient diary) in patients with rheumatoid arthritis: A comparison between personal digital assistant and paper-pencil format. Arthritis Rheum 2007 Apr 15;57(3):454-460 [FREE Full text] [doi: 10.1002/art.22613] [Medline: 17394232]
21. Rossi MC, Nicolucci A, Di Bartolo P, Bruttomesso D, Girelli A, Ampudia FJ, et al. Diabetes Interactive Diary. A new telemedicine system enabling flexible diet and insulin therapy while improving quality of life: An open-label, international, multicenter, randomized study. Diabetes Care 2010 Jan;33(1):109-115 [FREE Full text] [doi: 10.2337/dc09-1327] [Medline: 19808926]
22. Heron KE, Smyth JM. Ecological momentary interventions: Incorporating mobile technology into psychosocial and health behaviour treatments. Br J Health Psychol 2010 Feb;15(Pt 1):1-39 [FREE Full text] [doi: 10.1348/135910709X466063] [Medline: 19646331]
23. Miron-Shatz T, Stone A, Kahneman D. Memories of yesterday's emotions: Does the valence of experience affect the memory-experience gap? Emotion 2009 Dec;9(6):885-891. [doi: 10.1037/a0017823] [Medline: 20001131]
24. Diamantidis CJ, Zuckerman M, Fink W, Hu P, Yang S, Fink JC. Usability of a CKD educational website targeted to patients and their family members. Clin J Am Soc Nephrol 2012 Oct;7(10):1553-1560 [FREE Full text] [doi: 10.2215/CJN.03690412] [Medline: 22798537]
25. Diamantidis CJ, Zuckerman M, Fink W, Aggarwal S, Prakash D, Fink JC. Usability testing and acceptance of an electronic medication inquiry system for CKD patients. Am J Kidney Dis 2013 Apr;61(4):644-646. [doi: 10.1053/j.ajkd.2012.09.016] [Medline: 23157933]

Abbreviations

CKD: chronic kidney disease
eGFR: estimated glomerular filtration rate
ER: emergency room
FDA: US Food and Drug Administration
GEE: generalized estimating equation
GFR: glomerular filtration rate
IQR: interquartile range
IVDRS: interactive voice-inquiry dial-response system

MD: medical doctor
PRO: patient-reported outcome
SF-12: 12-item Short Form Health Survey
SF-36: 36-item Short Form Health Survey
SKC: Safe Kidney Care

Edited by G Eysenbach; submitted 05.10.15; peer-reviewed by K Flott, D Tscholl; comments to author 24.12.15; revised version received 20.01.16; accepted 28.01.16; published 26.05.16.

Please cite as:
Fink JC, Doerfler RM, Yoffe MR, Diamantidis CJ, Blumenthal JB, Siddiqui T, Gardner JF, Snitker S, Zhan M
Patient-Reported Safety Events in Chronic Kidney Disease Recorded With an Interactive Voice-Inquiry Dial-Response System: Monthly Report Analysis
J Med Internet Res 2016;18(5):e125
URL: http://www.jmir.org/2016/5/e125/
doi:10.2196/jmir.5203
PMID:27230267

Original Paper

Applying Multiple Methods to Comprehensively Evaluate a Patient Portal's Effectiveness to Convey Information to Patients

Jordan M Alpert[1], PhD; Alex H Krist[2], MPH, MD; Rebecca A Aycock[2], PhD; Gary L Kreps[3], PhD

[1]Department of Health Behavior and Policy, Virginia Commonwealth University, Richmond, VA, United States

[2]Department of Family Medicine and Population Health, Virginia Commonwealth University, Richmond, VA, United States

[3]Department of Communication, George Mason University, Fairfax, VA, United States

Corresponding Author:
Jordan M Alpert, PhD
Department of Health Behavior and Policy
Virginia Commonwealth University
830 E Main St
Richmond, VA, 23219
United States
Phone: 1 8046284790
Fax: 1 8046281233
Email: jordan.alpert@vcuhealth.org

Abstract

Background: Patient portals have yet to achieve their full potential for enhancing health communication and improving health outcomes. Although the Patient Protection and Affordable Care Act in the United States mandates the utilization of patient portals, and usage continues to rise, their impact has not been as profound as anticipated.

Objective: The objective of our case study was to evaluate how well portals convey information to patients. To demonstrate how multiple methodologies could be used to evaluate and improve the design of patient-centered portals, we conducted an in-depth evaluation of an exemplar patient-centered portal designed to promote preventive care to consumers.

Methods: We used 31 critical incident patient interviews, 2 clinician focus groups, and a thematic content analysis to understand patients' and clinicians' perspectives, as well as theoretical understandings of the portal's use.

Results: We gathered over 140 critical incidents, 71.8% (102/142) negative and 28.2% (40/142) positive. Positive incident categories were (1) instant medical information access, (2) clear health information, and (3) patient vigilance. Negative incident categories were (1) standardized content, (2) desire for direct communication, (3) website functionality, and (4) difficulty interpreting laboratory data. Thematic analysis of the portal's immediacy resulted in high scores in the attributes enhances understanding (18/23, 78%), personalization (18/24, 75%), and motivates behavior (17/24, 71%), but low levels of interactivity (7/24, 29%) and engagement (2/24, 8%). Two overarching themes emerged to guide portal refinements: (1) communication can be improved with directness and interactivity and (2) perceived personalization must be greater to engage patients.

Conclusions: Results suggest that simple modifications, such as increased interactivity and personalized messages, can make portals customized, robust, easily accessible, and trusted information sources.

(J Med Internet Res 2016;18(5):e112) doi:10.2196/jmir.5451

KEYWORDS

eHealth; patient portal; health communication; qualitative study; case study

Introduction

There is significant progress to be made in creating digital communication platforms that improve health outcomes [1]. While patient portals have been lauded as a method to enhance health communication [2], and they have been found to be helpful, their impact has not been as profound as anticipated. Yet patients are still enthusiastic about the portals' capabilities to assist in managing their health [3]. In one study, 43% of patients believed that apps, such as portals, could improve relationships with doctors, 48% would feel more in control of their health, and 40% would be would encouraged to ask physicians more questions [4]. Evidence has demonstrated that portals contribute to improved health outcomes by increasing

cancer screenings, especially when information was explained in lay language, used personalized recommendations, and provided educational resources [5]. Still, clinicians report dissatisfaction [6] and patients are underwhelmed with the design and functionality of portals [7].

Incorporating immediacy into portal design has the ability to increase their usability and importance as health communication resources. Immediacy refers to features that promote physical and emotional closeness, caring relationships, authenticity, and enthusiasm [8]. The construct has a rich tradition focusing on face-to-face communication in education [9,10], psychological counseling [11,12], and health care delivery [13,14]. However, immediacy has only recently been applied to digital health communication. Incorporating immediacy into the design of health communication tools can better engage, involve, and motivate patients to promote their health and well-being [15]. However, thus far there has been little evidence to demonstrate that portals can create a sense of immediacy [16].

Our study explored how a unique set of methods—critical incident reports from patients, focus groups of clinicians, and thematic analysis—can be used to evaluate and better inform the design of patient portals. We analyzed one exemplar patient portal, MyPreventiveCare [17], as a test of concept. MyPreventiveCare, a highly functional, prevention-focused online resource, was designed specifically to activate and engage patients around preventive care [18]. It is used by 12 practices in Virginia in the United States and reaches over 82,000 patients, but is being extended to an additional 300 practices in 15 states through a series of grants. In addition to providing laboratory results and viewing information from the medical record, the innovative portal customizes content based on hundreds of demographic, historical, behavioral, and clinical variables to make personalized recommendations and provide educational material based on current guidelines using content from HealthFinder.gov, a credible online health information resource from the National Health Information Center (Washington, DC). The portal provides information about needed cancer screenings, how to better monitor chronic conditions, and how to improve health behaviors. Nearly half (43.4%) of patients registered on MyPreventiveCare have logged in to the portal in the past year and frequently visit the following areas: laboratory results, medical record information, and preventive care recommendations. The average patient accesses the portal almost 4 times per year.

Methods

To sufficiently assess the test portal, we used multiple methods to collect and analyze data. Using multiple methodologies in health science research helps researchers view problems from multiple perspectives to enrich the meaning of a singular viewpoint [19] and contributes to developing a more complete understanding of a problem [20]. We describe the methodology for conducting interviews, focus groups, and the thematic analysis.

Interviews

Sample

Age demographics for current users of the portal are as follows: 18–24 years (3980/72,362, 5.50%), 25–34 (13,387/72,362, 18.50%), 35–44 (16,281/72,362, 22.50%), 45–54 (16,643/72,362, 23.00%), 55–64 (12,302/72,362, 17.00%) and ≥65 (9769/72,362, 13.50%). More females (41,246/72,362, 57.00%) use the system than males (31,116/72,362, 43.00%) and in fact, the most frequent users are women 45–54 years old. Although women 45–54 years old is the largest demographic group of users, criteria for interviews included patients ranging from 18–79 years, to generate as many viewpoints as possible.

We used naturalistic intercept sampling techniques [21], such as purposeful sampling [22] and convenience sampling [21], to recruit active users, age 18-79 years, who accessed the system at least once in the past year. To enhance recruitment, efforts were made to make participation in the interviews convenient for patients who already were at the clinic waiting for their doctors' appointments. Patients who met eligibility and had an upcoming appointment were contacted via telephone 1 week in advance to schedule an interview before the appointment with their physician. In addition, the medical staff identified eligible patients and referred them either before or after the appointment with their physician. We recruited participants in the fall of 2014 and winter of 2015 in 2 primary care practices that belong to the Virginia Ambulatory Care Outcomes Research Network. The study was described to patients in a private conference room, and then informed consent was reviewed. Scheduling interviews with patients around impending appointments over the course of several weeks allowed for a diverse cross-section of patients to be recruited.

Critical incident technique (CIT) [23] was the main method of inquiry, which is a qualitative research method with strong exploratory and investigative abilities [24-26]. The technique was designed as a flexible set of principles with the goal of identifying incidents that the users considered positive or negative and to probe participants about their experiences. The CIT is especially advantageous to the evaluation of a patient portal because it allows users to reflect on the most meaningful events from their experience using the system. Unlike usability studies, which solely focus on the way an individual operates a system, CIT allows both patients and clinicians to identify instances in which the portal affected their lives outside of directly using the system. CIT has been used to analyze quality of care [25] and is a common method in the application of health care services [27].

Procedures

During interviews, participants were asked to recall their best and worst experiences using the portal. These 2 broad questions framed the discussion, allowing the interviewer to focus on the incidents mentioned and investigate their significance. Interviews ceased and saturation was achieved when keywords and phrases were frequently repeated, participants' discourse was similar, and recurrent meanings were discovered [28]. All discussions were audio recorded and subsequently transcribed. The average interview time was 14 minutes in length.

Analytical Process

To ensure objectivity of the incidents collected, we developed a set of specific rules that specified whether an incident was positive or negative based on (1) the situation of the incident, (2) relevance to the general aims of the study, and (3) importance of the incident's effect on the aims [23]. Data analysis was guided by the tradition of CIT and the constant comparative method [29], in which 2 researchers (JA, GK) independently read each transcript, exploring for prevalent themes. Initially, any incident that could be considered either positive or negative was collected. Upon further analysis, the researchers discussed examples of positive and negative incidents, thereby refining the qualifications to be considered an incident worthy of inclusion. Next, the researchers shared observed themes and formed categories by creating an aggregated codebook. If discrepancies occurred, the researchers referred to the rules and aims of the study and discussed their perspective until consensus was met. Each researcher once again independently analyzed the data and assigned a code that best captured the significance of the incident. The researchers met to review their findings frequently and discussed any discrepancies that arose until they reached a consensus around a particular theme. To assess the validity of the data, we initiated respondent validation, or member checks [30], with several participants to confirm the interpretations of the findings.

Focus Groups

Sample

We conducted 2 focus groups with clinicians in 1 Northern Virginia medical office. The administrator of the portal emailed providers working in the medical office and made an announcement at a monthly meeting about participation in the study. Participation was voluntary and scheduled during providers' lunch hour. The first focus group included 8 physicians (2 men and 6 women), comprising 4 residents and 4 full-time physicians. The second focus group consisted of 5 participants (4 women and 1 man), 4 nurses and 1 emergency medical technician. The average length of both focus group sessions was 51 minutes.

Procedures

We used the CIT to allow clinicians to concentrate on the extreme positive and negative functions of the portal. A printed list of general questions was presented at the beginning of the focus group session. Questions were grouped into the following domains: positive experiences, negative experiences, and how the portal positively or adversely affected the communication process with patients and staff. The questions were semistructured, to allow for organic, flexible conversations that stimulated discussion [22]. Each focus group session was audio recorded and transcribed. For transcripts involving the focus group data, we used the same analytical process previously described for interviews.

Thematic Analysis of the Portal

We conducted the third method, thematic analysis, after interviews and focus groups were completed. We used patients' and clinicians' perspectives as a lens to analyze the content found on the portal, as well as by using key factors of immediacy identified in previous research [15,31,32]: user engagement, personalization, interactivity, enhances understanding, and motivates behavior change (definitions can be found in Table 1 [33-40]). We conducted the thematic analysis by reviewing the portal, searching for specific patterns related to immediacy. To strengthen claims of the thematic analysis, we incorporated numerical results, which enabled the amount of evidence in the data to be quantified to support the conclusions [41].

Table 1. Immediacy definitions for thematic analysis of the patient portal MyPreventiveCare

Immediacy feature	Definition
Interactivity	Definitions of interactivity typically focus on 2 measures: (1) 2-way flow of information and (2) rapid exchange of information. Other definitions include control as the main component of interactivity [33], meaning that participants should be able to exercise control over the communication exchange. For the purposes of this study, interactivity includes all 3 components and is considered to be reciprocal and synchronous communication that offers active control [34].
Personalization	According to e-commerce websites, personalization is "the adjustment and modification of all aspects of a website that are displayed to a user in order to match users' needs and wants" [35]. Although the portal is not an e-commerce website, the previous definition was thoroughly explicated and is relevant to this study.
Engagement	Combining definitions from the fields of marketing [36], media [37], and psychology [38], engagement is considered the level of an individual's physical, cognitive, and emotional involvement or connection with a specific medium.
Motivates behavior	Motivators are triggers, prompts, cues, or calls to action that encourage a user to take action [39].
Enhances understanding	Factors that enhance understanding, similar to relational understanding, in which an individual knows both what to do and why [40].

Analytical Process

The first author (JA) logged in to the portal as a test user and analyzed 27 separate content pages, of which 24 were evaluated. The home page, "Dashboard," and "Library" pages were not analyzed because they only functioned as navigational pages. Every webpage was examined individually to determine the prevalence of each immediacy characteristic. A checklist with each immediacy feature's definition was designed and individual webpages were meticulously scrutinized using the checklist. After an initial review, each page was checked a second time to confirm initial findings. After the first author conducted the analysis, each of the authors gave input in regard to the identification of themes. In particular, the portal's administrator reviewed the findings and provided feedback on actual patient and provider experiences. The first author then reevaluated each individual webpage taking into account the administrator's

perspective, as well as the theoretical framework of immediacy features.

The use of multiple methods (interviews, focus groups, and thematic analysis) and recruitment of key stakeholders (patients and clinicians) provided triangulation by using methods with different strengths and limitations to support a single conclusion [42].

This study was given full institutional review board approval by the George Mason University Office of Research Integrity & Assurance. Participants signed an informed consent document and, to ensure confidentiality, audio files were only accessed by the researcher, and personal information, such as names, were de-identified upon transcription.

Results

The study included 44 total participants in 31 patient interviews (18 women and 13 men) and 2 focus groups (13 clinicians). We collected a total of 142 incidents, 102 negative and 40 positive.

Patient Interviews

Incidents classified as positive (31/113, 27.4%) were outweighed by negative incidents (82/113, 72.6%). The following themes were most salient throughout positive and negative incidents cited by patients.

Positive Incidents

Patient interviews revealed 3 main categories of the portal's usefulness: (1) the ability to instantly access medical information, (2) availability of clear health information, and (3) patient vigilance.

Instant Access

Nearly half of all positive incidents were associated with the instantaneous retrieval of medical information. For example, patient #1, a busy mother, anxiously awaited test results but was unable to call the office during business hours. She appreciated how she could log in to the portal in the evening and look up the results.

Similarly, most patients complained that obtaining laboratory results was a chore involving endless telephone calls, but the portal streamlined the once-laborious process. Patient #8 described how he used to get results before the portal was available. He said, "The office doesn't have time to call you, or they call, and they don't get you and leave a message, and you call back and leave a message."

Clear Health Information

Not only did patients appreciate the ability to get laboratory results, but also the portal's design made it easy to navigate. Upon logging in, patients were presented with a dashboard featuring large, colorful icons used as the website's primary navigation. Referring to the portal's interface, patient #30 commented, "It is very icon-driven and it's all very intuitive, so I don't find it complicated to use." In addition, throughout every section of the website, a dictionary is available to look up challenging terminology.

Patient Vigilance

The ease of the portal's navigation contributed to patients being able to carefully monitor their health. Patients used information found on the portal to gauge their health status. Patient #4 monitored her laboratory results to determine whether "medications were working or if there was a problem somewhere." Correspondingly, patient #11 always checked laboratory results as soon as they were available and said, "If my labs are getting worse, or even if they're normal, I can make adjustments and monitor it closer." The availability of health information, such as laboratory results, allowed patients to check differences from previous tests, while prevention recommendations, like getting the flu shot, reinforced efficacious behavior and served as a reminder.

Negative Incidents

Negative incidents were classified into 4 main categories: (1) standardized content, (2) the desire for clinicians to communicate directly with the patient, (3) website functionality, and (4) difficulty interpreting laboratory data.

Standardized Content

Many patients were unaware that content was personalized based on their electronic medical record and self-reported data. For instance, patient #3 said, "it felt generic. I don't know if it was specific to my medical history, but it didn't feel personal." Patient #12 agreed and said, "I'd want something more focused on me as opposed to something generic like this."

Desire for Direct Communication

After reading the portal's content, many patients craved information directly from the clinician. For instance, patient #15 wondered, "Can the doctor and the medical provider put what they would propose as next steps?" Although the content that appeared was generated from updates the clinician made to the patient's record, the language used did not reflect that the content was personally delivered by a clinician. This sentiment was best expressed by a retired woman, patient #10, who was inspired with new questions and concerns after viewing content from the portal. She often consulted advice nurses, but since they were no longer available, she yearned for the portal to be a substitute.

In addition to the desire to interact directly with clinicians, many patients wanted a resource that would be beneficial between scheduled office visits. Patient #4 stated, "once in a while something goofy will happen, so it would be nice to come in here (portal) and read the things that I need." Similarly, patient #2's blood pressure often wildly fluctuates and during those times, she thought, "it would be nice if I could say, is this something I should worry about?"

Website Functionality

The portal was susceptible to problems experienced by many websites, such as issues logging in and server crashes. Indeed, 20% (16/82) of all negative patient incidents involved an instance in which the portal did not function correctly. Patient #12 remembered a time when he had difficulty recovering his password. He said, "I couldn't remember what my password

was. Even after I got the password, when I went to log in again, it wouldn't let me."

When patients were able to log in, data entry mistakes or database errors sometimes caused the portal to inaccurately report results. Patient #5 said, "It's been hit or miss as far as having the information being correct." Also, information was not updated to appropriately reflect their health status. Patient #28 cited, "I kept getting these flash things that said I was overdue on a pap exam. I had just been in in April, but it kept flashing at me like I was a bad girl."

Laboratory Data Interpretation is Difficult

Patients appreciated the portal's preventive care messages, such as a recommendation to get a mammography, but experienced difficulty interpreting diagnostic laboratory information. This accounted for 11% (9/82) of negative incidents. For instance, patient #15 referred to the "Watch Your Weight" page and said, "One question I have is this BMI [body mass index]. Is it just sort of a number pulled out of the air or is it in fact appropriately calculated?" He wanted to better understand what the number actually meant to "take the number seriously." Similarly, some patients had trouble deciphering the results of the blood sugar section, as exemplified by patient #27, who examined the chart on the webpage and wondered, "It was measured on September 3rd. I have a value of 100 and my goal is to be less than 126, so I should be good? I don't know what that symbol means. [Orange icon indicating a marginal score]. Why would I be marginal?"

Clinician Focus Groups

We combined data from the 2 focus groups, one with physicians and another with rooming staff. Clinician incidents were comparable with patient data, with 69% (20/29) negative and 31% (9/29) positive.

Positive Incidents

Positive incidents among clinicians were classified into 2 main areas: (1) patients feel empowered (5/9, 56%) and (2) the portal can generate office efficiency (4/9, 44%).

Develops Patient Empowerment

Over half of all positive incidents involved the belief that patient access to medical information was beneficial. A nurse recalled, "I work with patients who have jobs that take them out of the country and the fact that they have this information is awesome." Before the introduction of the portal, nurses would review laboratory results over the phone, but patients did not have access to the data as a reference. Furthermore, clinicians observed higher levels of patient motivation after interacting with the portal. A nurse said, "[Patients] will use the portal to track things, like 'my cholesterol wasn't that good, so I need to increase my exercise', and next time, they'll come in and say, 'I added another day to my exercise routine and I'm really anxious to see how my cholesterol is now.'"

Due to increased levels of motivation, patients experienced more productive office visits. A physician noticed that the portal "helps start a conversation when there's so much to cover." Another physician confirmed that notion and said the portal created "patient led agendas rather than the physician leading the agenda." He continued, "I've had patients who have already looked at their laboratory results before they come in." Physicians noticed that patients altered their behavior and even researched possible treatment options before their appointment.

Generates Workplace Efficiencies

Although many more negative incidents focused on how the portal has the potential to create additional work, positive incidents highlighted how the portal contributes to a more efficient working environment. For instance, reminders delivered to patients via the portal lessened the need for unit clerks to call and remind patients about upcoming examinations. A physician agreed that reminders were particularly useful because they prevented unit clerks from interpreting medical information. In addition, less time was devoted to playing "phone tag" because patients could correct errors, such as 1 incident in which a patient's file stated that her last colonoscopy was 12 years ago, even though the procedure was performed 5 years ago.

Negative Incidents

Clinicians' incidents overlapped with patients', and in many cases, reinforced what patients experienced. Negative incidents fell into 3 main categories: (1) lack of feedback, (2) fears that the portal can increase workload, and (3) inappropriate use of the system.

Lack of Feedback

Clinicians considered the portal a valuable tool, but acknowledged that it contained flaws. Perhaps the biggest flaw, which accounted for 40% (8/20) of negative incidents, was that clinicians could not confirm whether patients viewed or understood information that was input into the portal. A nurse summarized the problem by describing how she typed the analysis of laboratory results or medication directions, but was unaware whether patients either saw the information input or understood its meaning.

In other scenarios, clinicians received inquiries from patients, but the portal was not deemed an appropriate forum to discuss the matter. For example, a resident recalled the following predicament: "You get this message and you want to answer it...but I find that if that question pushes the boundaries of what I should do outside of an office visit, maybe I should bring them in? But they're asking for it, so I send the anti-biotic I otherwise wouldn't."

Another issue that compressed communication between patients and clinicians was the belief that the portal was not equipped to handle complex communication. A doctor brought up the issue of cancer screening and said, "The problem is there are different ways to do it. We all know we should screen for colon cancer. When I see a patient in person, I ask them do you want to do a colonoscopy or annual stool test?" According to clinicians, a much richer communication platform was needed to conduct meaningful conversations.

Fears of Increased Workload

Although clinicians cited specific instances in which the portal generated efficiency in the office, there was still speculation that increased use of the portal would create additional work. Nurses were worried that patients would call the office at greater

frequencies because sometimes "patients look at laboratory results and one thing is a little off, so they freak out and they're calling us and asking." Physicians were also concerned about devoting more time to addressing patients' questions because, as one doctor said, "It is frankly time we don't get reimbursed for." Since physicians did not receive confirmation that the patient viewed the message, they had to take the time to call the patient and ensure the information was communicated.

Inappropriate Use of the System

Surprisingly, clinicians acknowledged that they sometimes purposefully contributed complex medical jargon. For instance, nurses said that younger doctors entered sophisticated terms because they wanted to impress their supervisors who may see the information they input. Physicians with more seniority were also at fault. Nurses complained that many doctors wrote messages meant for nurses and were unaware that patients were also capable of viewing them.

Sometimes, physicians recognized that content on the portal was not written in a sensitive manner. A doctor said, "I've had a few patients who get the health maintenance reminders and say, 'my portal says I'm fat.'" Other doctors felt that the content on the portal may be suitable for the average patient, but might be beyond comprehension for patients who are not native English speakers.

Thematic Analysis of the Portal

On the basis of attributes of immediacy (interactivity, personalization, engagement, motivates behavior, enhances understanding), we examined the portal as to whether the characteristics were present (summarized in Table 2).

Table 2. Presence (+) or absence (–) of immediacy characteristics in the patient portal MyPreventiveCare.

	Enhances understanding	Personalization	Motivates behavior	Interactivity	Engagement
Preventive care recommendations					
Summary page	–	+	+	+	+
Take aspirin	+	+	+	–	–
Get tested for diabetes	+	+	+	–	–
You have high blood pressure	+	+	+	–	–
Get a tetanus shot	+	+	+	–	–
Other recommended behavior					
Mammogram	+	+	+	+	–
Cervical cancer	+	+	+	+	–
Colon cancer testing	+	+	+	+	–
Cholesterol	+	+	+	–	–
Diet	+	+	+	–	–
Exercise	+	+	+	–	–
Smoking	+	+	–	–	–
Weight	–	+	–	–	–
Pneumonia vaccine	+	+	+	–	–
Flu vaccine	+	+	–	–	–
Bone density	+	+	+	–	–
Laboratory results					
Your labs	–	+	–	–	–
Edit information section					
Update page	+	+	+	+	+
Library					
Dictionary	–	–	–	–	–
Setting priorities	–	–	–	–	–
Prevention topics A–Z	+	–	+	+	–
Self-management tools	+	–	+	+	–
Help					
Help page	+	–	+	–	–
Contact us					
Contact page	n/a[a]	–	–	–	–
Characteristics present, n (%)	19/23 (78%)	18/24 (75%)	17/24 (71%)	7/24 (29%)	2/24 (8%)

[a]n/a: not applicable.

Enhances Understanding

This category measured whether content allowed patients to understand not only what to do, but why to do it. We did not review the "Contact" page because its content was not applicable. This section received the highest score, with 78% (18/23) of the portal's content contributing to enhancing understanding.

The way in which information was presented was integral to enhancing patients' understanding. Patient #28 commented how the bulleted lists made it easier to process recommendations. Moreover, the use of graphics supported the text it accompanied. For example, an image on the "Exercise" page of a man playing basketball contributed to reinforcing healthy behavior, while photographs of individuals eating fresh vegetables on the "Diet" page emphasized the concept of healthy eating and made it more appealing. Showcasing positive behaviors supplemented the information presented and served as a model to apply recommendations. Other factors that enhanced understanding were detailed explanations. For instance, on the "Take aspirin"

page, an initial header outlined the patient's health status, such as whether they smoked and whether they were taking an aspirin dosage. Next, a brief explanation on the benefits of aspirin appeared, followed by the criteria for whether aspirin is appropriate, next steps, and several related links to more information about aspirin, as well as heart disease. Similarly, T-scores were described in the "Bone density" section, including how weight influences blood pressure. Throughout the portal, simple and direct language was used, demonstrated by the average word length at only 6 letters, which contributes to enhancing patient understanding.

Personalization

Personalization was determined by an enhanced sense of inclusion and cooperation to match users' needs and wants. The main way personalization was achieved came through language, such as using pronouns ("we" or "you"), as well as providing unique information applicable to an individual's life. For instance, recommendations were offered, like "You are due to get another tetanus shot now" and, after listing a patient's medical information, the following text was generated: "These conditions place you at higher risk for heart disease." The portal's customized content contributed to 75% (18/24) of pages having the quality of personalization.

Although we gave credit for personalization, terms like "you" and "your" were inconsistent. On the "Your labs" page, information was written in a very clinical manner, negating past attempts at personalization. The text read, "Patient would like a mammogram," instead of using "you" or the patient's name. This was especially true on the dashboard, which was the first page that appeared when a patient logged in. At the top, it read, "You are here," instead of immediately establishing personalization by using the patient's name.

Other areas of the portal lacking personalization were the "Dictionary" and "Prevention topics" pages, which listed topics that could be irrelevant to patients, such as information about cervical cancer to male patients. Lastly, stock photographs of physicians were used throughout the portal instead of photos of clinicians familiar to patients.

Motivates Behavior

Aspects of the portal that motivated positive health behavior had to be actionable, meaning participation was encouraged to perform the activity, like scheduling an exam. Over 70% (17/24, 71%) of content had features that could motivate behavior. Every page within the "Preventive care you need now" section fulfilled the motivational behavior criteria by using verbs that promoted action, like "talk with your doctor" and "protect yourself." Understanding that losing weight is challenging, on the "Your next steps" section of the "Diet" page, the text attempted to motivate patients by stating, "Keep eating five or more servings of vegetables and fruits per day." However, we did not consider the "Smoking," "Weight," and "Flu vaccine" pages to be motivational because the content was overly passive. Confirming this analysis, patient #31, with diabetes, thought that the content should be "scarier" to motivate people to change their behavior. She acknowledged that losing weight would assist with her diabetes management, but after reviewing the content, she said, "I'm not scared enough to be doing all the things I ought to be doing." Similarly, the "Flu" page recommended as a next step, "Talk with your doctor about getting a flu shot." Although an action verb was used and previous content provided an overview of benefits, no urgency to get the vaccine was generated.

Interactivity

Interactivity included synchronous communication, the ability of a user to take control, or any function that enabled the user to become involved more deeply within the content. We found only 29% (7/24) of the pages to have characteristics of interactivity. There were very few opportunities to interact with the portal aside from reading text. Interactive quizzes, cholesterol calculators, and other risk assessment tools were presented as external links, which took the user away from the portal. Lack of interactivity was confirmed by a doctor who was frustrated over the system's asynchronous communication. He said, "The patient can't contact us directly. Instead, they have to go through the nurse and we review it a few hours later, or a day later, and then it becomes broken communication." However, a relatively strong aspect of the portal's interactivity was the ability for the user to control the navigation and choose the area accessed. For instance, the "Self-management tools" page allowed for patients to find specific information relevant to their condition, while the mammogram and cervical cancer pages enabled the patient to input dates of their last test and whether abnormal results appeared.

Opportunities for greater interactivity were available on most pages. For instance, the "Take aspirin" page asked questions about risk, but there was no way for the patient to answer those questions and receive timely feedback. Also, the "Your labs" page had a row of physician comments, but it was not possible for the patient to respond, make an appointment, or get more information.

Engagement

Engagement received the lowest score with 8% (2/24). Only 2 webpages were deemed engaging, having content that made the user want to further explore the website by creating physical, cognitive, or emotional involvement.

The main "Preventive care you need now" page made bold proclamations, like "You have high blood pressure," which would get a patient's attention and impel them to further explore. The rest of the pages did not have enough personalized content to be engaging. For example, the "Mammogram" page made a recommendation for women age 40–84 years instead of providing more specific advice based on the patient's specific age. The "Self-management tools" page included links within 19 specific content categories, including exercise, healthy diet, and smoking cessation, but required that a patient scroll through almost 200 links.

Consequences of Incidents and Immediacy Levels

We synthesized the patients' and clinicians' incidents with results of the content analysis to identify domains that would improve website functionality. We discovered the following 2 themes: (1) communication can be improved with directness and interactivity and (2) standardization contributes to patient disengagement.

Communication Can Be Improved With Directness and Interactivity

While the intention of the portal was to provide direct information to augment the patient's care, both patients and clinicians found that the portal sometimes hampered the communication process. The portal provided general awareness overviews; however, patients yearned for specific information about how a particular treatment personally affected them. Therefore, patients used other online sources for more specific information.

Patient #9 said, "Sometimes I have back pain and I cannot come to the primary care doctor, but I need some information." This patient's comments aligned with many other patients' views, whose first instinct was to seek outside sources rather than rely on the portal. Furthermore, patients needed detailed information between visits to manage chronic diseases. For instance, patient #31, with type 2 diabetes, was concerned about managing her blood sugar. She said, "I'm not sure I'm always doing the right things" and the portal could not be counted on as a resource to help her manage specificities of the disease. Patients became disinterested in the portal's content, because as patient #19 said, "I pretty much know all this stuff." Patient #21 considered the portal "dumbed down" and mentioned that, since he was due for a prostate exam, he already conducted basic research. Perhaps, if patients deemed content to be authoritative and felt that it was coming directly from the clinician, they would find it more helpful and seek out the portal between scheduled visits. By using alternative websites, patients would combine the vetted and personalized information from the portal with recommendations from the other websites, which may or may not be beneficial given the patient's specific medical history.

Similarly, clinicians reported that they were unable to effectively communicate in an appropriate manner through the portal. They were unsure whether patients read their instructions, and asynchronous communication patterns disrupted care.

Standardization Contributes to Patient Disengagement

Although content was personalized, patients largely viewed the portal as just another platform offering standardized information, which minimized patients' assessment of it as a resource for delivering personalized health recommendations.

Lack of Personal Relevance

Patient #25 became indifferent and dismissed the portal as a helpful tool after reading the section about weight loss. He responded, "Yeah, but what do I do with it? I'm overweight. I need to lose weight. Got it." Patient #8 was unimpressed with the recommendations and said, "A lot of the information, the basics, this is stuff I could read on CNN health's website. [It] doesn't tell me much." Similarly, patient #20 was disappointed with the portal's lack of guidance about vitamins. Since the efficacy of vitamins is not evidence based and therefore would not be recommended by the clinician, the portal's content did not address vitamins.

Not Motivational

Overall, the portal did not provide actionable information, leaving patients with little motivation to use the system. Patient #27 said, "This is not motivating. Maybe if it gave me examples, like go for a walk or have an apple with lunch." Lack of specificity was best exemplified by patient #4, who visited the physician for a flu shot even though it had been years since she last received the vaccine. She thought the portal's flu shot content was "generic" and it did not inspire her, but said, "I am here today because I've been told by my daughter's baby's pediatrician that I need to get a flu shot."

Discussion

Principal Findings

Conducting a multiple methodological study on a patient portal as a case study revealed that many factors need to be considered when delivering personalized health information via a portal. It was necessary to use multiple methods and involve numerous stakeholders to truly understand the complexity of the user experience. The combination of gathering critical incidents from patients and clinicians, and analyzing the themes of the portal's content enabled us to simultaneously confirm the results through a concurrent triangulation strategy [43].

Findings suggest that personalized content must contain higher levels of engagement and interactivity. This may be accomplished through increased levels of immediacy, which has been associated with greater patient satisfaction, understanding [13,44], and compliance [45]. Communication strategies incorporating immediacy, like adding personalized content addressing patients' concerns about treatment, can contribute to humanizing interactions between patients and clinicians, encouraging patient participation and building trust [8].

Making portals more immediate through other forms of health information technology hold the opportunity for greater levels of engagement [46]. For instance, interactive videos to inform patients about elective back surgery helped facilitate decision making and informed consent [47]. Compared with reading text, watching videos has the ability to improve patients' knowledge of health risks [48]. Furthermore, videos featuring information directly from a patient's clinician could make the content seem more authoritative and personal.

Although implementation and integration of health portals has initially been slow [49], dramatic growth is expected due to patient demand and physician adoption [50]. Therefore, it is necessary to design patient-centric systems that provide comprehensive, coordinated value to patients, while also meeting the needs of clinicians [51]. In fact, clinician endorsement and engagement with the portal is an important factor that influences patient adoption [52].

The portal used for this case study has been at the forefront of personalizing content, but even more is needed. While the majority of webpages fulfilled the definition of personalization, patients expected a higher level of personalization akin to sophisticated mobile phones and social networking websites. This tension between patients' expectations of personalization and clinician guidelines is pronounced because patients' expectations of the portal's main function differ from clinicians'. While the goal of the portal used in this study was primarily to

educate and motivate patients for preventive screenings, patients considered the portal to be a mechanism to organize all of their health care needs. Sections with the highest level of personalization were on the preventive care pages, but patients expected similar personalization levels on laboratory results and medication pages. Based on this study's findings, the objective of portals needs to be better communicated to patients, and future modifications should consider patients' expectations. A portal rich with immediacy not only has the capacity to fulfill patients' desires, but also has the potential to improve the medical environment. Results from this study indicated that clinicians worry about increases in workload, but they also realize the potential for portals to create greater efficiency.

Limitations and Future Directions

There were several limitations to this study. It would have been beneficial to have greater representation of adults from various age ranges, ethnicities, and socioeconomic classes. Participants represented only a small subset of all users, and recruitment likely selectively targeted those with strongly positive or negative impressions. In addition, comparisons across other health care information systems would have been beneficial to contrast differences and illuminate the various ways that systems are used. Future research should focus on analyzing the content, design, and usability of portals. In addition, research should analyze how patients use the portal's content to determine whether it is a factor influencing health behavior.

Conclusions

Through the analysis of this portal, we found that higher levels of immediacy are necessary to sufficiently motivate patients to take preventive care measures. Portals are becoming a fixture of every medical office; therefore, it is necessary to modernize portals so that they can achieve their potential of enhancing health communication and improving health outcomes. The methods used in this study can be replicated when analyzing other patient portals, and conclusions can inform designers to develop more effective health information communication.

Acknowledgments

This research was supported by grants from the National Cancer Institute (R01CA166375-01A1) and the National Center for Advancing Translational Sciences (UL1TR000058).

Conflicts of Interest

Virginia Commonwealth University holds the intellectual property rights to the interactive preventive care record evaluated in this study. Although the university and developers are entitled to the system's revenue, MyPreventiveCare is a noncommercial product, and no revenues have been generated other than grant funding.

References

1. Kreps GL. Achieving the promise of digital health information systems. J Public Health Res 2014 Dec 2;3(3):471 [FREE Full text] [doi: 10.4081/jphr.2014.471] [Medline: 25553317]
2. Pagliari C, Detmer D, Singleton P. Potential of electronic personal health records. BMJ 2007 Aug 18;335(7615):330-333 [FREE Full text] [doi: 10.1136/bmj.39279.482963.AD] [Medline: 17703042]
3. Tieu L, Sarkar U, Schillinger D, Ralston JD, Ratanawongsa N, Pasick R, et al. Barriers and facilitators to online portal use among patients and caregivers in a safety net health care system: a qualitative study. J Med Internet Res 2015;17(12):e275 [FREE Full text] [doi: 10.2196/jmir.4847] [Medline: 26681155]
4. Undem T. Consumers and Health Information Technology: A National Survey. Oakland, CA: California HealthCare Foundation; 2010 Apr 13. URL: http://www.chcf.org/~/media/MEDIA%20LIBRARY%20Files/PDF/PDF%20C/PDF%20ConsumersHealthInfoTechnologyNationalSurvey.pdf [accessed 2016-04-29] [WebCite Cache ID 6h8HnnSSN]
5. Krist AH, Woolf SH, Rothemich SF, Johnson RE, Peele JE, Cunningham TD, et al. Interactive preventive health record to enhance delivery of recommended care: a randomized trial. Ann Fam Med 2012;10(4):312-319 [FREE Full text] [doi: 10.1370/afm.1383] [Medline: 22778119]
6. Edsall RL, Adler KG. User satisfaction with EHRs: report of a survey of 422 family physicians. Fam Pract Manag 2008 Feb;15(2):25-32 [FREE Full text] [Medline: 18303652]
7. Lazard A, Watkins I, Mackert MS, Xie B, Stephens KK, Shalev H. Design simplicity influences patient portal use: the role of aesthetic evaluations for technology acceptance. J Am Med Inform Assoc 2016 Apr;23(e1):e157-e161. [doi: 10.1093/jamia/ocv174] [Medline: 26635314]
8. Kreps GL. Engaging health communication. In: Socha TJ, Pitts MJ, editors. The Positive Side of Interpersonal Communication. New York, NY: Routledge; 2012:249-258.
9. Frymier A. A model of immediacy in the classroom. Commun Q 1994 Mar;42(2):133-144. [doi: 10.1080/01463379409369922]
10. Henning Z. From barnyards to learning communities: student perceptions of teachers' immediacy behaviors. Qual Res Rep Commun 2012 Jan;13(1):37-43. [doi: 10.1080/17459435.2012.719207]
11. Egan G. Interpersonal Living: A Skills-Contract Approach to Human-Relations Training in Groups. Monterey, CA: Brooks/Cole Pub Co; 1976.

12. Lee LA, Sbarra DA, Mason AE, Law RW. Attachment anxiety, verbal immediacy, and blood pressure: results from a laboratory-analogue study following marital separation. Pers Relatationships 2011 Jun;18(2):285-301 [FREE Full text] [doi: 10.1111/j.1475-6811.2011.01360.x] [Medline: 21647240]
13. Richmond V, Smith RS, Heisel AD, McCroskey JC. Nonverbal immediacy in the physician/patient relationship. Commun Res Rep 2001 Jun;18(3):211-216. [doi: 10.1080/08824090109384800]
14. Liu Y, Malin JL, Diamant AL, Thind A, Maly RC. Adherence to adjuvant hormone therapy in low-income women with breast cancer: the role of provider-patient communication. Breast Cancer Res Treat 2013 Feb;137(3):829-836 [FREE Full text] [doi: 10.1007/s10549-012-2387-8] [Medline: 23263740]
15. Kreps GL, Neuhauser L. Artificial intelligence and immediacy: designing health communication to personally engage consumers and providers. Patient Educ Couns 2013 Aug;92(2):205-210. [doi: 10.1016/j.pec.2013.04.014] [Medline: 23683341]
16. Greenhalgh T, Hinder S, Stramer K, Bratan T, Russell J. Adoption, non-adoption, and abandonment of a personal electronic health record: case study of HealthSpace. BMJ 2010;341:c5814 [FREE Full text] [Medline: 21081595]
17. MyPreventiveCare: Helping You Take Care of Your Health. Richmond, VA: VCU Department of Family Medicine; 2016. URL: https://secure.mypreventivecare.com/ [accessed 2016-04-29] [WebCite Cache ID 6h8KLDoSe]
18. Krist A, Woolf S, Rothemich S, Longo D, Johnson R, Matzke G, et al. An Interactive Preventive Health Record (IPHR) to Promote Patient-Centered Preventive Care. 2011 Feb 28. URL: https://healthit.ahrq.gov/sites/default/files/docs/publication/r18hs017046-krist-final-report-2011.pdf [accessed 2016-04-27] [WebCite Cache ID 6h5ARtVLO]
19. Creswell JW, Klassen AC, Plano Clark VL, Smith KC. Best Practices for Mixed Methods Research in the Health Sciences. Bethesda, MD: Office of Behavioral and Social Sciences Research; 2011 Aug. URL: https://obssr-archive.od.nih.gov/mixed_methods_research/pdf/Best_Practices_for_Mixed_Methods_Research.pdf [accessed 2016-04-27] [WebCite Cache ID 6h5AlvrV5]
20. Clark V, Creswell JW, Plano CVL. Understanding Research: A Consumer's Guide. Upper Saddle River, NJ: Merrill; 2010.
21. Marshall MN. Sampling for qualitative research. Fam Pract 1996 Dec;13(6):522-525 [FREE Full text] [Medline: 9023528]
22. Tracy SJ. Qualitative Research Methods: Collecting Evidence, Crafting Analysis, Communicating Impact. Hoboken, NJ: Wiley-Blackwell; 2013.
23. Flanagan JC. The critical incident technique. Psychol Bull 1954 Jul;51(4):327-358. [Medline: 13177800]
24. Woolsey LK. The critical incident technique: an innovative qualitative method of research. Can J Counselling 1986 Oct;20(4):242-254.
25. Kemppainen JK. The critical incident technique and nursing care quality research. J Adv Nurs 2000 Nov;32(5):1264-1271. [Medline: 11115012]
26. Chell E. Critical incident technique. In: Cassell C, Symon G, editors. Qualitative Methods and Analysis in Organizational Research: A Practical Guide. London, UK: Sage Publications; 2004:45-60.
27. Beech P, Norman IJ. Patients' perceptions of the quality of psychiatric nursing care: findings from a small-scale descriptive study. J Clin Nurs 1995 Mar;4(2):117-123. [Medline: 7704381]
28. Owen W. Interpretive themes in relational communication. Q J Speech 2009 Jun 05;70(3):274-287. [doi: 10.1080/00335638409383697]
29. Charmaz K. Constructing Grounded Theory: A Practical Guide Through Qualitative Analysis. London, UK: SAGE; 2006.
30. Lincoln YS, Guba EG. Naturalistic Inquiry. Beverly Hills, CA: Sage Publications; 1985.
31. Bartlett Ellis RJ, Carmon AF, Pike C. A review of immediacy and implications for provider-patient relationships to support medication management. Patient Prefer Adherence 2016;10:9-18 [FREE Full text] [doi: 10.2147/PPA.S95163] [Medline: 26792985]
32. Hill CE, Gelso CJ, Chui H, Spangler PT, Hummel A, Huang T, et al. To be or not to be immediate with clients: the use and perceived effects of immediacy in psychodynamic/interpersonal psychotherapy. Psychother Res 2014;24(3):299-315. [doi: 10.1080/10503307.2013.812262] [Medline: 23848833]
33. Rogers EM. Diffusion of Innovations. New York, NY: Free Press; 2003.
34. Liu Y. Developing a scale to measure the interactivity of websites. J Advertising Res 2003 Aug 13;43(02):207-216.
35. Wu D, Im II, Tremaine M, Instone K, Turoff M. A framework for classifying personalization scheme used on e-commerce websites. In: Proceedings of the 36th Annual Hawaii International Conference on System Sciences. 2003 Presented at: Annual Hawaii International Conference on System Sciences; 2003; Big Island, HI.
36. Patterson P, Yu T, Ruyter KD. Understanding customer engagement in services. In: Advancing Theory, Maintaining Relevance. 2006 Presented at: ANZMAC 2006 conference; 2006; Brisbane, Australia.
37. Calder BJ, Malthouse EC. Media engagement and advertising effectiveness. Kellogg Advertising Media 2008:1-36. [doi: 10.1002/9781119198154.ch1]
38. Higgins ET. Value from hedonic experience and engagement. Psychol Rev 2006 Jul;113(3):439-460. [doi: 10.1037/0033-295X.113.3.439] [Medline: 16802877]
39. Fogg BJ. A behavior model for persuasive design. In: Proceedings of the 4th International Conference on Persuasive Technology. 2009 Presented at: International Conference on Persuasive Technology; 2009; Claremont, CA.
40. Skemp RR. Relational understanding and instrumental understanding. Arithmetic Teacher 1978 Nov;26(3):9-15.

41. Maxwell JA. Qualitative Research Design: An Interactive Approach (Applied Social Research Methods). Los Angeles, CA: Sage Publications; 2012.
42. Fielding NG, Fielding JL. Linking Data: The Articulation of Qualitative and Quantitative Methods in Social Research. Beverly Hills, CA: Sage Publications; 1986:41-53.
43. Creswell JW. Research Design: Qualitative, Quantitative, and Mixed Methods Approaches. Thousand Oaks, CA: Sage Publications; 2009.
44. Larsen KM, Smith CK. Assessment of nonverbal communication in the patient-physician interview. J Fam Pract 1981 Mar;12(3):481-488. [Medline: 7462949]
45. Burgoon J, Pfau M, Parrott R, Birk T, Coker R, Burgoon M. Relational communication, satisfaction, compliance - gaining strategies, and compliance in communication between physicians and patients. Commun Monogr 1987 Sep;54(3):307-324. [doi: 10.1080/03637758709390235]
46. Ahern DK, Woods SS, Lightowler MC, Finley SW, Houston TK. Promise of and potential for patient-facing technologies to enable meaningful use. Am J Prev Med 2011 May;40(5 Suppl 2):S162-S172. [doi: 10.1016/j.amepre.2011.01.005] [Medline: 21521591]
47. Deyo RA, Cherkin DC, Weinstein J, Howe J, Ciol M, Mulley AG. Involving patients in clinical decisions: impact of an interactive video program on use of back surgery. Med Care 2000 Sep;38(9):959-969. [Medline: 10982117]
48. Steffenino G, Viada E, Marengo B, Canale R, Nursing and the Medical Staff of the Cardiac Catheterization Unit. Effectiveness of video-based patient information before percutaneous cardiac interventions. J Cardiovasc Med (Hagerstown) 2007 May;8(5):348-353. [doi: 10.2459/01.JCM.0000268131.64598.49] [Medline: 17443101]
49. Rhodes HB. The PHR quandary: despite the benefits, issues of technology and trust slow adoption. J AHIMA 2007 Apr;78(4):66-7, 69. [Medline: 17455850]
50. EMR 2015: The Market for Electronic Medical Records. New York, NY: Kalorama Information; May 2015.
51. Baird A, Nowak S. Why primary care practices should become digital health information hubs for their patients. BMC Fam Pract 2014;15:190 [FREE Full text] [doi: 10.1186/s12875-014-0190-9] [Medline: 25421733]
52. Irizarry T, DeVito Dabbs A, Curran CR. Patient portals and patient engagement: a state of the science review. J Med Internet Res 2015;17(6):e148 [FREE Full text] [doi: 10.2196/jmir.4255] [Medline: 26104044]

Abbreviations

CIT: critical incident technique

Edited by G Eysenbach; submitted 17.12.15; peer-reviewed by S Woods, T Irizarry; comments to author 03.04.16; revised version received 18.04.16; accepted 18.04.16; published 17.05.16.

Please cite as:
Alpert JM, Krist AH, Aycock RA, Kreps GL
Applying Multiple Methods to Comprehensively Evaluate a Patient Portal's Effectiveness to Convey Information to Patients
J Med Internet Res 2016;18(5):e112
URL: http://www.jmir.org/2016/5/e112/
doi:10.2196/jmir.5451
PMID:27188953

Original Paper

Therapeutic Affordances of Online Support Group Use in Women With Endometriosis

Amie Shoebotham[1], MSc; Neil S Coulson[1], PhD

Division of Rehabilitation & Ageing, School of Medicine, University of Nottingham, Nottingham, United Kingdom

Corresponding Author:
Neil S Coulson, PhD
Division of Rehabilitation & Ageing
School of Medicine
University of Nottingham
Medical School, Queen's Medical Centre
Nottingham, NG7 2UH
United Kingdom
Phone: 44 1158466642
Fax: 44 1158230231
Email: neil.coulson@nottingham.ac.uk

Abstract

Background: The Internet has provided women living with endometriosis new opportunities to seek support online. Online support groups may provide a range of therapeutic affordances that may benefit these women.

Objective: To examine the presence of therapeutic affordances as perceived by women who use endometriosis online support groups.

Methods: Sixty-nine women (aged 19-50 years, mean 34.2 years; 65.2% (45/69) United Kingdom, 21.7% (15/69) United States) participated in a Web-based interview exploring online support group use. Participants had been using online support groups for an average of 2 years and 4 months (range = 1 month to 14 years, 9 months). Responses were analyzed using inductive thematic analysis.

Results: The analysis revealed 4 therapeutic affordances related to online support group use: (1) "connection," that is, the ability to connect in order to support each other, exchange advice, and to try to overcome feelings of loneliness; (2) "exploration," that is, the ability to look for information, learn, and bolster their knowledge; (3) "narration," that is, the ability to share their experiences, as well as read about the experiences of others; and (4) "self-presentation," that is, the ability to manage how they present themselves online. The associated outcomes of use were predominantly positive, such as reassurance and improved coping. However, a number of negative aspects were revealed including the following: concerns about the accuracy of information, arguments between members, overreliance on the group, becoming upset by negative experiences or good news items, and confidentiality of personal information.

Conclusions: Our findings support the previously proposed SCENA (Self-presentation, Connection, Exploration, Narration, and Adaptation) model and reveal a range of positive aspects that may benefit members, particularly in relation to reassurance and coping. However, negative aspects need to be addressed to maximize the potential benefit of support groups. Some of these can be addressed relatively easily through making privacy policies clearer, including health professionals to moderate content, and structuring forums to encourage the sharing of positive stories.

(J Med Internet Res 2016;18(5):e109) doi:10.2196/jmir.5548

KEYWORDS

endometriosis, social network, social support, qualitative research

Introduction

Endometriosis is a chronic condition that affects women of reproductive age. It is defined as the presence of endometrial tissue outside the uterine cavity, which induces a local inflammatory response [1]. Common symptoms include chronic pelvic pain, dysmenorrhea, fatigue, heavy menstrual bleeding, and dyspareunia. Estimates of the prevalence of the condition range between 2% and 17% of the female population [2];

however, the prevalence can increase to up to 47% in women with fertility issues [3]. Women typically experience a delay before diagnosis [4-6] with long-term care focused on symptom management [7,8]. Psychosocial effects are also common, including high levels of depression and anxiety [9-11], emotional distress [12,13], and poorer quality of life [9,14]. Endometriosis can negatively affect social life, work, and daily activities [15-17] as well as education, finances, life opportunities [18], personal relationship quality [8], and physical intimacy [19].

Women living with endometriosis typically need considerable support to meet the various challenges associated with the condition [20], and access to it may yield benefits through a reduction in illness-related stress [21] and improvements in well-being [22]. However, women with endometriosis often describe a lack of support [23,24], encounter feelings of social isolation [13], and find it difficult to disclose their condition and symptoms to others because of embarrassment or a fear of not being believed [16]. It can be difficult for these women to find an understanding and knowledgeable source of support, especially as the support may be needed for many years [20]. Finding other women with endometriosis, who understand the constant struggles, is therefore potentially a valuable source of support for them [16,20]. However, women often lack contacts with other women with the condition [23], or they may be unable to attend local support groups, should any exist, because of pain restrictions. The Internet therefore becomes a key resource when seeking the support of other women living with endometriosis [16,20].

Over recent years, the popularity of using the Internet for information and support has been growing; in 2013, 43% of adults in the United Kingdom [25] and 72% in the United States [26] accessed health information online. Furthermore, up to a quarter of Internet users have gone online to read about other people's experience of a health-related problem and 16% to find others with the same problem [26]. The popularity of using the Internet for such purposes may be explained by its flexibility; it does not have the same restrictions as face-to-face support groups with regard to set times or locations [27]. The lack of face-to-face contact can also help users feel less inhibited and talk openly about embarrassing topics [28] or stigmatized health conditions [29].

Previous research has suggested that women with endometriosis do find the Internet to be a valuable source of information regarding both the condition and its medical management [30,31]. To date, however, there is lack of research into the use of online support groups for endometriosis. In contrast, research exploring the role of online support groups across other long-term conditions has been extensive and has found their use to have benefits including the provision of emotional support [32], instilling hope [33], fostering empowerment [32,34], and reducing feelings of isolation [35].

However, although such benefits are common, there can also be negative aspects of online support group use. Research has indicated that their use can lead to distress, particularly when reading about others' negative experiences [35]. Users can also become obsessed with the groups [35] or display disinhibition due to the anonymous interface of the groups [28]. Furthermore, there is a risk that the medical information that is shared online is potentially inaccurate [36].

In order to understand how the use of online support groups might influence the outcomes of those living with endometriosis, this study considered the therapeutic affordances of online support group use. The origins of affordance theory stem from perceptual and cognitive psychology and are based on how individuals perceive the objects in their environment, pertaining to both what the object is and what potential uses it affords [37]. The properties of an object will therefore contribute to its perceived affordance. Furthermore, the differing experiences, beliefs, and goals of an individual will lead to different perceived affordances, and so affordances are unique to each individual [38]. The emphasis in the theory is on the interaction between the user and the object and its resulting outcomes.

Therapeutic affordances can therefore be seen as the "actionable possibilities" of the object as determined by the user [39], and the object in this instance is online support groups. Studying therapeutic affordances allows us to look at both the uses and effects of online support groups. This research will therefore provide further insight into the use of online support groups than previous research, as it will examine the underlying mechanisms that drive the observed effects and outcomes of support group use [39].

Previous research on therapeutic affordances within a health context is lacking. However, Merolli et al [39] examined the therapeutic affordances of social media use in people with chronic pain. They identified five therapeutic affordances of social media: self-presentation, connection, exploration, narration, and adaptation. The results highlight both the "actionable possibilities" of social media use, that is, connection, and the therapeutic value of such, that is, not feeling alone. Their work provides a good point of comparison for our study; however, there are important differences such as the health condition being studied and the different "object." Online support groups have different properties than social media and therefore may enable different affordances. Specifically, online support groups tend to use an asynchronous discussion forum format, through which individual group members can post messages to start new conversations, or reply to messages within existing conversations, thereby creating a tree-like structure of discussion threads.

The aim of this study therefore was to explore the therapeutic affordances of online support group use in women with endometriosis and the resulting outcomes. The results of this research will yield a richer understanding of both the therapeutic benefits of online support groups and the needs of women living with endometriosis.

Methods

Procedure

This research involved the recruitment of women (aged 16 years or older) who use online support groups for endometriosis. They were recruited through messages posted onto the boards of such support groups, during a 4-week recruitment period from June 16 to July 13, 2015. In order to identify potential recruitment

channels, searches were conducted on Google, using terms such as "online support group endometriosis" and "endometriosis support group," as well as similar searches on Facebook. The groups that appeared to be the most recently active (some had been dormant for a number of months) were selected as potential recruitment channels.

Initial contact was made with the board moderators to seek permission to recruit their group members via a message posted onto the boards of the groups. It was made clear to the moderators that the research focused on women's experiences of using such groups, and they were given the opportunity to review the message before posting. Once permission was granted, the recruitment message was posted, which explained the aims of the study and invited those interested to click on a link to the survey, hosted by Bristol Online Surveys. There was a degree of variation in the final boards used for recruitment, having 1 general endometriosis online support group, 1 board hosted by an online fertility support group, and 1 group hosted by Facebook.

When directed to the survey page, participants were provided with additional information about the study, and they were then required to complete a consent form before beginning the survey. Participants first completed a series of short answer questions relating to their background (date of birth, country of residence, diagnostic status, and so on) and use of online support groups (frequency, period of use). Next, they were invited to respond to a set of open-ended questions that explored their motives and experiences of using online support groups and whether their use has any effect on how they cope with or manage the condition.

Data analysis

In order to provide a rich account of the data, the responses to the open-ended questions were qualitatively analyzed using deductive-inductive semantic thematic analysis [40]. This analysis approach allows for the identification of common themes across the dataset and therefore closely reflects the data and the language used by participants. This is particularly important given the aims of this study.

The analysis was carried out by the first author (AS) according to the guidelines proposed by Braun and Clarke [40], and QSR's NVivo 10 software was used to maintain an audit trail. First, to become familiarized with the data, the transcripts were read and reread several times and initial ideas about the data were noted down. Next, interesting features of the data were coded, keeping codes close to the language used by participants where possible. Codes were then organized into meaningful groups to form potential themes. Codes that appeared consistently throughout the data were automatically considered as a potential inductive theme. Data relevant to each potential theme were gathered, and then themes were reviewed, refined, and given a clear definition and name. Reviewing the common language used in these themes allowed them to be organized into the final therapeutic affordances. To reduce bias and check whether the themes reflected the data, an independent researcher read through some of the transcripts and agreement was reached on the final themes.

Ethical Considerations

Before data collection, the research was reviewed and approved by the University of Nottingham (United Kingdom) institutional ethics review committee. In order to maintain ethical requirements, participants were provided with full study information that informed them of their rights as a research participant, procedures of withdrawal, and full contact details of the researchers. Before beginning the survey, participants were required to complete a Web-based consent form indicating that they understood such rights. They were each asked to provide a unique password to allow for their data to be withdrawn if they so wished after survey completion and were given a time frame within which they could withdraw. They were informed that, to maintain privacy and confidentiality, no personally identifiable information would be reported in the results. Furthermore, they were able to indicate during consent if they were happy for their responses to be quoted in research reports (3 participants chose not to have their responses quoted). In order to meet ethical requirements, quotes provided in the report only identify participants by their responder number (ie, Respondent 17).

Results

In total, 69 participants completed the Web-based survey. The respondents were aged between 19 and 50 years, with a mean age of 34.2 years, and the majority were residents of the United Kingdom (65.2% 45/69) or the United States (21.7% 15/69). Of the respondents, 66 (95.7%) had received a confirmed diagnosis of endometriosis, which was made between 1 month and 20 years before survey completion (mean time since diagnosis = 4 years, 1 month). Participants had been using online support groups for endometriosis for between 1 month and 14 years, 9 months (mean use period = 2 years, 4 months), and most participants (59.4% 41/69) accessed the groups more than 5 days a week.

Following deductive-inductive thematic analysis, the therapeutic affordances that were present in the data were "connection," "exploration," "narration," and "self-presentation." Quotes were largely positive; however, participants did refer to some negative aspects afforded by online support groups. Table 1 describes the inductive themes that emerged from the data analysis within each affordance and the key language used by participants.

Table 1. Final therapeutic affordances, inductive themes, and descriptive language.

Therapeutic affordance	Themes	Language
Connection	Contact with others with endometriosis	Talk
	Support	Advice
	Exchanging advice	Alone
	Forming relationships	Support
	Mitigating isolation	Friends
	Freedom of access	Community
		Understanding
Exploration	Information seeking	Find
	Credibility	Learn
	Learning/knowledge	Discover
	Empowerment	Look
		Know
Narration	Sharing experience	Stories
	Reassurance	Share
	Disheartening	Hear
		Experiences
		Competition
		Negativity
Self-presentation	Privacy	Confidential
	Disclosure	Private
		Embarrassing

Connection

The most frequently cited affordance of online support groups was the ability to connect with others, and participants used this connection to support each other, exchange advice, and to try to overcome feelings of loneliness. They benefited from the freedom to connect with a large number of individuals and at all hours of the day or night.

Participants frequently commented on their ability to connect with other women who have endometriosis through the online support groups. They demonstrated a sense of relief at finding those who "understand" what they are going though: "...to be able to connect with others suffering the same just gives you some piece of mind" [Respondent 28] and "...having a place to speak freely with people who completely understand & offer love, encouragement & advice is a great help at times of need" [Respondent 11]. Finding an understanding support network had also significantly helped women to cope with the condition: "After years of struggling with GP's finding other people who understood played a huge part in making me feel normal again" [Respondent 25] and "I think one of the reasons I haven't had a nervous break down or gotten depressed is because of the groups. I have people who understand what I'm going through and can give me advice" [Respondent 10].

Participants spoke positively of the support they were able to offer each other through online support groups: "The support is great" [Respondent 29] and "[I] have found the support online invaluable" [Respondent 4]. In some instances, this support had a marked effect on women's lives, enabling improved coping and confidence in seeking better treatments: "...talking through some of the darker times with women who know what I'm going through has been hugely helpful and I feel I am coping better than I was before" [Respondent 16], "I also received support from lots of ladies when I didn't know how to make myself heard..." [Respondent 32], and "I am also going to ask to be referred to a specialist because of support from others on this forum" [Respondent 5].

Receiving support was described by many as a key motivation for online support group use. However, some women also talked favorably about providing support to others, finding the process to be rewarding, and even of psychological benefit: "I feel I can provide some help & support to the ladies because I have a lot of experience..." [Respondent 11] and "...the rewarding feeling following directly supporting others has improved my mental wellbeing..." [Respondent 16].

A large number of participants spoke about how online support groups had allowed them to share advice, either on general matters, or those related to treatments or self-management behaviors. Participants often used words such as "appreciate" or "valuable" when talking about advice received from others: "The advice and support received is invaluable..." [Respondent 25], "I also really appreciate the advice and support these groups offer" [Respondent 27], and "I have had valuable advice on symptom relief, side effects of treatments, post op care etc etc" [Respondent 51]. Women also found it helpful to provide advice to others, experiencing a positive feeling from doing so: "...I feel like it is helpful to me to be able to help other women, with comments and advice" [Respondent 13].

Advice in general was gratefully received, although some participants expressed caution at the medical advice shared through these platforms: "...I'm not going to take medical advice

through Facebook" [Respondent 42]. Some indicated that although they do read the advice, they would usually ultimately check it with their medical professionals and follow their advice: "I feel that the online group helps answer any questions...but I have always gone ahead with what I thought or the medical staff say" [Respondent 33] and "I do listen to the advice but in the end I usually go with what my consultant suggests even if it's not what I want" [Respondent 67].

The most frequently occurring theme within "connection" was that of mitigating isolation; the online support groups had allowed participants not to feel alone. The word "alone" was mentioned by several respondents, and many talked about the isolating nature of endometriosis: "I feel less alone in the groups- I know I'm not the only one experiencing this" [Respondent 22], "Also knowing you are not alone in your struggle, as endometriosis can be very isolating..." [Respondent 9], and "I was feeling so alone and very down before I discovered there was others just like me who understood what a struggle constant pain could be" [Respondent 25].

Having found others in a similar position to themselves, women were able to feel acceptance of their condition: "Now I know that there are other women suffering I can accept that this is actually happening to me!" [Respondent 22] and "...knowing that what I'm feeling is real...that I'm not alone" [Respondent 23]. They were also able to feel reassured about any unusual symptoms experienced: "There are symptoms I've experienced that I thought was unique to me only to find that many others have the same issues. It was through the online support groups that I found I wasn't alone" [Respondent 60].

Connections made through online support groups could be very strong, with some talking of the relationships they had formed with other people. They often used words such as "friends" and "community" when talking about their connections with others: "I've made quite a few friends who I speak with regularly" [Respondent 22], "I've made good friends" [Respondent 20], and "...I have increasingly used the forum to seek community..." [Respondent 8]. One woman particularly emphasized the role of these online support groups in enabling connections she could not have found elsewhere: "I find a sense of community there that I can't find in the physical world" [Respondent 8].

However, connecting to others through the support groups did not always foster positive outcomes. Participants alluded to problems in communication that they had witnessed between other members, such as arguments: "There are sometimes arguments about what is the best treatment" [Respondent 22] and "I have seen many posts get out of order & things can escalate quickly..." [Respondent 11]. This was also reiterated by a member of the administration team for one support group: "As an admin on [name of OSG], I do sometimes have to deal with arguments and the usual sort of online drama..." [Respondent 50]. Although such incidents caused frustration, they were only mentioned by a minority of participants. One participant even found the positive side to it, citing it as a source of entertainment: "I visit the group now far more due to fear of missing a scandal! It's a hive of bitchiness - better than a soap opera" [Respondent 18].

The final theme relating to "connection" was how online support groups allowed participants the freedom to access people. This could be access to a large number of people: "In the groups I could ask 100s /1000s of people questions..." [Respondent 22]; or access at any time of the day or night: "The main benefits of being in an online support group is that you have access 24/7" [Respondent 1]. Access to such high volumes of people meant that online support groups enabled women to get help and support quite quickly if they needed it: "Any time I have a question I can post and get reassurance and answers rather quickly" [Respondent 22]. This could be particularly comforting at times when participants were suffering physically or psychologically; "...I have had three laparoscopies in the past and to know others are there whenever in need is lovely" [Respondent 29] and "Knowing that there is always someone there to listen when things get bad or I am feeling down" [Respondent 67].

A small number of women described potential pitfalls arising from this freedom of access. Having constant access to such platforms could cause preoccupation: "Possibly a bit pre occupied with the group" [Respondent 41] and "My life revolves round Facebook etc because I can do so little" [Respondent 18]. Furthermore, one woman felt that access to so many people actually negated the support she was seeking: "No one responds to my posts. There are too many posts/people and not enough support to go around" [Respondent 24].

Exploration

The second largest affordance of online support groups for those with endometriosis was "exploration." Women were able to look for information, learn, and bolster their knowledge, and in doing so appeared to feel more empowered.

Participants frequently described how their use of online support groups had enabled them to seek out information, from general information about the condition to more specific information on treatments, research, and specialists. Many commented on how they were able to find information that had not been provided by medical professionals: "Additional information that may not have been given by my own consultant" [Respondent 44] and "I do not think I would have obtained the high quality of this information from the medical community directly" [Respondent 37]. Again, positive words were used when describing the information gathered through online support groups: "...obtaining invaluable information on the condition..." [Respondent 14] and "...and overall wealth of information has been immensely helpful" [Respondent 38].

Although most participants valued the information found through this platform, some expressed concerns about the quality of information being given on such sites: "However it was / is also a minefield of myths and misinformation..." [Respondent 61] and "I feel like someone a bit younger might fall pray to some of the misinformation some women provide" [Respondent 2]. In a similar vein, some questioned the credibility of the sources of information and were concerned about people self-appointing themselves as experts: "There are some who think they are experts but really are not" [Respondent 51]. Participants generally felt the information was credible if it came from a medical expert: "I tend to only really listen to the

advice from the specialists who post and answer questions, regarding current research" [Respondent 55]. Alternatively, a number of women stated that they will do further research on any information obtained: "I conduct my own web based research in order to validate information I have obtained from online support groups" [Respondent 37].

A common theme talked about by participants was how online support groups had facilitated their learning and knowledge. This learning encompassed a wide range of factors, as it could be about the disease itself, the symptoms, treatment options, or self-management techniques: "In the forums I've learned so much about what the disease is, the proper surgery (excision vs ablation), the best doctors, and how to deal with the day to day misery of this disease" [Respondent 62] and "...I've learned about books to read, and diet, and supplements, and various other ways to deal with the symptoms" [Respondent 20]. Again, participants alluded to how the learning facilitated by these groups went above and beyond what they could have received from their doctors: "The information received from the medical community doesn't even come close to the details of what I have learned in online Support groups" [Respondent 37].

This new knowledge received through the online support groups facilitated improvements in the self-management of the condition: "It has improved my health literacy around managing my own condition, which I am really grateful for" [Respondent 8] and "I know about staying ahead of the pain now. In times where I have previously gone to hospital with extreme pain I have piggy-backed painkillers as suggested in forums and been able to manage things at home" [Respondent 19]. It also helped to foster a sense of reassurance for the women, that what they have been experiencing is normal: "I also understand a lot more so I can rationalise what is happening to me" [Respondent 22] and "It has helped me understand why I have continued pain, despite multiple surgeries, and what I need to do to hopefully help revolve [*sic*] or at least reduce it" [Respondent 50].

A particularly pertinent outcome of the learning and knowledge facilitated by online support groups was a sense of empowerment. This was mentioned by many participants: "Knowing more about the condition from the group has made me more empowered" [Respondent 19] and "I feel a lot more empowered now that I have a reasonable understanding of the disease, this allows me to feel more in control and be better informed in the decisions I make about treatment" [Respondent 61]. In fact, many participants talked about how they now felt in control or more knowledgeable about their treatment options. As a consequence, they were able to attend medical appointments feeling more prepared and confident: "...the information obtained from online support groups helps me to be more prepared when speaking with medical practitioners..." [Respondent 37] and "I feel more confident about seeing medical professionals and getting a good standard of care as the support groups have provided me with sources of information that I have used to educate myself" [Respondent 19].

Narration

Although slightly modest in size compared with the previous two affordances, participants did frequently comment on their use of online support groups to share their experiences, as well as hear the experiences of others. Through such narration, and that of others, participants were provided with a sense of reassurance, although at times it could also cause feelings of distress and disheartenment.

The majority of the "narration" quotes related to accessing the experiences of others, but a few did talk about how they had shared their own experiences with the condition. In doing so, participants described positive feelings and reassurance: "My mental health has improved by being able to share my experience..." [Respondent 16] and "I had a few really painful examinations I shared this with the group and found support and was also told this was not uncommon with woman with endo..." [Respondent 68]. Some also mentioned the benefits of sharing experiences as a two-way process: "I also benefit from the fact that we are able to discuss and share our experiences and outcomes with specialists..." [Respondent 50] and "It is nice to share stories and hear from others about their journey with endometriosis" [Respondent 3].

The responses that related to accessing the experiences of others were relatively evenly spread with regard to whether they fostered positive or negative outcomes. On a positive note, participants talked about how hearing those experiences have helped them: "Seeing other people's stories, be it success or negative stories as you can learn from both and gain some insight..." [Respondent 62] and "The personal experiences the other women have had and their outcomes of different treatments. Helps me to decide what to do next" [Respondent 43].

The experiences of others provided participants with feelings of reassurance about what they themselves had been experiencing. For example, one woman stated, "I feel reassured that many women with endo have similar frightening symptoms" [Respondent 7]. Another found that it helped to validate her own experience: "Hearing the testimonies of women living all over the world and knowing that what I'm feeling is real..." [Respondent 23]. Others simply found reassurance when reading others' experiences by realizing that things could be worse, for example: "it's made me realise how lucky I am that I manage very well day to day compared to a lot" [Respondent 4].

However, many women also found that accessing others' experiences could be disheartening. Some commented on the tendency for others to only share sad or negative stories: "There are so rarely any positive posts. People only seem to feel the need to post when they have something bad to report" [Respondent 18] and "The problem is that everyone talks about problems, very bad stories that make you feel more depressed, you never read happy things or happy stories!" [Respondent 42]. Hearing such negativity from others could be quite distressing: "Have felt disheartened and even a little depressed when reading some women's experiences..." [Respondent 49] and "...people repeating on multiple posts that it destroyed their lives to not have kids. As someone who has been diagnosed since 14 and is only 22 now it can be very overwhelming..." [Respondent 65].

At times participants described the propensity for some women's narratives to be almost competitive in nature, which was seen as counterproductive, as it prevented them from wanting to

share anything themselves: “I rarely comment in these groups now because it's like a competition. My endometriosis is worse than your endometriosis. I need to take more pills than you. My gynaecologist says mine is the worst case he's seen etc.” [Respondent 18]. One woman also talked about the opposite, how other women could be patronizing about the severity of others’ symptoms if they had not had them as badly themselves. The negative outcomes she felt were quite extreme: “There are also other ladies who tend to be a bit patronising and don't accept that endo can be different for everyone...This can make me feel quite insecure and pathetic when I am off work for the 6th time this year...” [Respondent 32].

Finally, although it was mentioned by many that there was a tendency for women to post negative stories, many participants reported that when positive stories were shared, such as a new pregnancy, it could cause a lot of distress in a group of people who frequently suffer from fertility issues. Although none of the women themselves said they felt this way, they often reported seeing other people getting upset about it, and sometimes even having arguments: “I have seen other members argue with each other about posting sensitive information, such as posting pregnancy announcements without ‘trigger warnings’...” [Respondent 37] and “There are those that, like myself, are infertile because of Endo. They cannot stand to see women make announcements that after having surgery or some form of treatment they are now pregnant...they will get upset and type angry rants and basically shoot down another EndoSister that is having a good experience right now” [Respondent 60].

Therefore, it would appear that rather than the potential positive outcomes of sharing such success stories such as feelings of hope, they may in fact exacerbate the distress felt by some women about their condition.

Self-Presentation

This affordance relates to how the users present themselves to others via online support groups. Although only a small number of participants alluded to this affordance, those that did placed great emphasis on its importance. To some, the confidential nature, or “privacy,” of the support groups was beneficial: “Knowing i can discuss something, in private, amongst fellow sufferers” [Respondent 58] and “I knew it was confidential because it was a closed group. Only people within the group could see- not all my facebook friends and family!” [Respondent 22]. Having this private, anonymous platform enabled some to disclose matters they might not usually, or allowed them to “preserve” their offline relationships: “It helps me deal with it ‘privately’ because I don't have to ‘bother’ the people I am close with and spend every day with about my health concerns. I just ask my endo sisters” [Respondent 13] and “Having an outlet for sharing personal and embarrassing symptoms...” [Respondent 16].

Conversely, some worried about how confidential the platform really was, stating fears over whether other online friends were able to see what they were posting: “One concern that I have is the fact that I am posting very sensitive medical and personal information about my symptoms onto Facebook, and I do not feel that my personal information is secure on this platform despite the fact that these online support groups are on ‘secret’ pages” [Respondent 37]. “The only problem I have is that not all the support groups are ‘secret’ or private. Though sometimes I like to share the details of this disease with others who experience the same, I don't particularly want everyone of my Facebook friends to know about all my period pains or symptoms such as infertility” [Respondent 39].

This issue appeared to be mainly apparent in those who were recruited from the online support group hosted by Facebook, on which users are not anonymous, and usually go by their own name rather than a pseudonym. However, their language used indicates it as an important factor in their use of online support groups.

Discussion

The aim of this research was to explore the therapeutic affordances of using online support groups for endometriosis. The results revealed 4 therapeutic affordances in this population: “connection,” “exploration,” “narration,” and “self-presentation.” In comparison with the results of previous research of online support groups, this study has uncovered similar therapeutic effects, including emotional support, reduced isolation, improved coping, reassurance, knowledge, and empowerment [32,34,35]. However, this research has also enabled a more in-depth exploration of the underlying mechanisms that generate the observed therapeutic outcomes, which suggests that the study of therapeutic affordances is a valuable way to research online support groups. Furthermore, the positive outcomes observed from online support group use have provided an indication of the potential influence that social support can have on those who are managing endometriosis.

The resulting therapeutic affordances are comparable to those found by Merolli et al [39] in their research on social media use in people with chronic pain. They too found the same 4 therapeutic affordances; however, their research also uncovered a fifth affordance, “adaptation.” This refers to the way that social media allows users to adapt their self-management needs in relation to their health status or illness flare-ups. In our research, only 2 participants provided quotes that could have been viewed as suggestive of this theme, but they were coded into “connection” and the “freedom of access” theme, as they talked of how they were able to access online support even at times of illness flare-up. Therefore, the SCENA model (Self-presentation, Connection, Exploration, Narration, and Adaptation) proposed by Merolli et al [39] to explain their results may be considered a valid model despite there being some differences, which can arguably be related to the different sample being examined and the fact that our work focused solely on online support groups rather than a range of social media.

The lack of evidence for the “adaptation” affordance in this research could simply be because it was not commented on by participants or because the affordance does not exist here. It is important to consider the differences in the “object” studied, as this research focused on online support groups rather than social media as studied by Merolli et al [39]. “Social media” encompasses a wider variety of platforms such as blogs, social networking sites, video-sharing sites, and chat rooms. This wider

variety could afford different uses and outcomes than online support groups alone. This might too account for the different emergent themes found within each affordance compared with Merolli et al [39]. Another possible explanation for the differing findings is that different health conditions were examined. Users perceptions may therefore differ based on their goals, experiences, and beliefs, which is an important factor in affordance theory [38].

Outside of the health domain, researchers have attempted to further develop the concept of therapeutic affordances by categorizing them into types based on their underlying function [41-43]. The results of this study can be interpreted using the framework proposed by Zhao et al [43], in which the perceived affordances are categorized into physical, cognitive, affective, and control functions. "Connection" may be seen as an affective affordance, allowing users to find a community, seek support, and feel less alone. "Exploration" may best be categorized as a cognitive affordance as it helps to facilitate the user's thinking and learning. "Self-presentation" may fall under control, as it allows users to control their environment, by maintaining their privacy for example. However, as Zhao et al [43] explain, affordances do not need to fall into these categories exclusively; for example, "narration" may be both affective, in providing reassurance or distress, and cognitive as it allows users to learn from others' experiences.

These theoretical frameworks allow us to understand the user's perceptions of and reactions to online support group use. By interpreting the results in relation to previous research in endometriosis, we can further improve our understanding of the affordances we have found. "Connection" was the most prominent therapeutic affordance in this study, with women frequently expressing their gratitude at finding others with endometriosis, who could understand their struggle and offer them support. Past research has reported that contact with other women with endometriosis is highly valued in this population [16,20]; however, it can be difficult to obtain [23]. This might explain the large number of responses relating to "connection," as using online support groups can enable them to achieve such contact. Many women in our sample talked of how alone and isolated their condition made them feel, which has been a common theme in previous qualitative research [12,13,31]. The lack of an understanding support network further intensifies this isolation [13], and so finding those with experience and knowledge can be a great sense of support [20].

The "exploration" affordance also received significant attention in this study, and in particular, women talked of the ability to use online support groups to search for information related to their treatment. Previous literature has highlighted the struggles that women face with regard to their treatment, including lack of effective treatments [13], pain continuing post treatment [23], and concerns over side effects [44]. Furthermore, women have often felt let down by medical professionals, or complained of a lack of information from them [4,12,13,18], feelings which were also reiterated in this study. They have therefore taken it upon themselves to search for the information they need, and the Internet and online support groups provide a platform to do this [31]. Finally, research has indicated the tendency for women to want to self-manage their condition, and the empowering experiences of such [16,45], which again was a common theme in this research.

Comparisons with previous research suggest that the affordances enabled by online support groups, which were revealed in this research, may be effective aids when coping with and managing endometriosis. These results imply that the use of online support groups might form a meaningful part in the self-management of endometriosis. However, until clinicians can make decisions about whether to recommend their use for such, further research is needed that accurately measures the effect of support groups on health outcomes. Recently, Merolli et al [46] studied patient-reported outcomes (PROs) of social media use in those with chronic pain and found that it had a positive effect on psychological, social, and cognitive aspects of health. Similar research is needed to corroborate these findings in those with endometriosis before online support group use can be recommended as a tool for self-management.

It is also necessary to consider the negative outcomes that were reported in the research, such as feeling disheartened by others' experiences. By accurately measuring PROs in future research, it might be possible to determine exactly how much of an effect these negative experiences are having on group users. It may also be useful in future research to try to measure the potential contribution of individual differences on perceived affordances. It may be the case that the therapeutic affordances offered by online support groups are only effective for certain types of individuals. Furthermore, our research cannot be generalized to other conditions, and so to further our understanding of the therapeutic affordances of online support group use, more research is needed to consider them in other chronic health conditions.

Our findings highlight a number of concerns that online support groups users may have that may affect the quality of their online experience. Our analyses revealed concerns over the accuracy and credibility of online information. One possible strategy to address these concerns might be to deploy an online moderator, such as a health professional, who can directly intervene to correct any misinformation or to validate information provided by a group member. Other "lay" moderators may similarly have a role to monitor online activity and ensure that members engage in a mutually respectful way. In terms of privacy concerns, a clearer statement of the privacy policy within online support groups may help to address this issue, although it should be noted that such concerns were restricted to Facebook groups. Finally, although there is clear value in people sharing their experiences, views, and opinions online, we did note that both negative and positive stories can have a detrimental effect on users. This is a difficult issue to address but may be dealt with partly through the development of specific subforums that focus on specific issues or experiences. In addition, there may be subforums devoted solely to "good news" and members would, we hope, have more freedom to choose the content they wish to be exposed to.

Study limitations

This study has a number of potential limitations. First, it is not possible to determine how representative this sample is of the whole population. Recruitment messages indicated that

participants would be asked about their experiences of using such groups, and so there may have been a tendency for only those with more positive experiences to participate. Second, although the recruitment message was posted on 3 online support groups, more than half of respondents (62.3% 43/69) were recruited from 1 group, the one hosted by Facebook. The support groups used varied considerably, both by platform and by group purpose; for example, one focused more on fertility issues associated with endometriosis, whereas another provided users with access to specialists. Although the motives for accessing the groups were fairly standard across respondents, these differences may have enabled users with different opportunities and therefore different affordances. Some evidence for this was highlighted by the "self-presentation" theme, in which participants from the Facebook group expressed more concerns over the privacy of the group than those recruited from the other 2 groups. Unfortunately, as specific between-group comparisons were not made, it is not possible to understand the potential influence of group type on the resulting affordances. However, the inclusion of a variety of group types will, we hope, have provided a more diverse range of participant characteristics and experiences.

Conclusion

This study has offered a unique exploration into the therapeutic affordances of online support group use in women with endometriosis. Through our analysis of the responses by participants to a series of open-ended questions, it has been possible to identify a number of therapeutic affordances arising from engagement with online support groups. Our findings reveal many positive aspects that may potentially benefit group members. However, several negative aspects are also revealed, particularly in relation to communication, personal information, and the presence of negative stories online. A number of strategies are recommended to mitigate these problems, including the presence of an explicit privacy policy, sections of the group where positive stories can be shared, and the presence of a health professional to address misleading or inaccurate information.

Acknowledgments

The authors would like to thank both the moderators of the online support groups used in this research for allowing recruitment of their members and all the participants for their valuable involvement in the research. We would also like to thank the independent researcher who assisted with the data analysis.

Conflicts of Interest

None declared

References

1. Kennedy S, Bergqvist A, Chapron C, D'Hooghe T, Dunselman G, Greb R, ESHRE Special Interest Group for EndometriosisEndometrium Guideline Development Group. ESHRE guideline for the diagnosis and treatment of endometriosis. Hum Reprod 2005 Oct;20(10):2698-2704 [FREE Full text] [doi: 10.1093/humrep/dei135] [Medline: 15980014]
2. Culley L, Law C, Hudson N, Denny E, Mitchell H, Baumgarten M, et al. The social and psychological impact of endometriosis on women's lives: a critical narrative review. Hum Reprod Update 2013;19(6):625-639 [FREE Full text] [doi: 10.1093/humupd/dmt027] [Medline: 23884896]
3. Meuleman C, Vandenabeele B, Fieuws S, Spiessens C, Timmerman D, D'Hooghe T. High prevalence of endometriosis in infertile women with normal ovulation and normospermic partners. Fertil Steril 2009 Jul;92(1):68-74. [doi: 10.1016/j.fertnstert.2008.04.056] [Medline: 18684448]
4. Ballard K, Lowton K, Wright J. What's the delay? A qualitative study of women's experiences of reaching a diagnosis of endometriosis. Fertil Steril 2006 Nov;86(5):1296-1301. [doi: 10.1016/j.fertnstert.2006.04.054] [Medline: 17070183]
5. Hadfield R, Mardon H, Barlow D, Kennedy S. Delay in the diagnosis of endometriosis: a survey of women from the USA and the UK. Hum Reprod 1996 Apr;11(4):878-880 [FREE Full text] [Medline: 8671344]
6. Husby GK, Haugen RS, Moen MH. Diagnostic delay in women with pain and endometriosis. Acta Obstet Gynecol Scand 2003 Jul;82(7):649-653. [Medline: 12790847]
7. Bergqvist A, Theorell T. Changes in quality of life after hormonal treatment of endometriosis. Acta Obstet Gynecol Scand 2001 Jul;80(7):628-637. [Medline: 11437721]
8. De GA, D'Hooghe T, Dunselman G, Dirksen C, Hummelshoj L, Simoens S, et al. The significant effect of endometriosis on physical, mental and social wellbeing: results from an international cross-sectional survey. Hum Reprod 2013;11:2677-2685. [doi: 10.1093/humrep/det284]
9. Sepulcri R, do AV. Depressive symptoms, anxiety,quality of life in women with pelvic endometriosis. Eur J Obstet Gyn R B Jan 31 2009;142(1):53-56. [doi: 10.1016/j.ejogrb.2008.09.003]
10. Low WY, Edelmann RJ, Sutton C. A psychological profile of endometriosis patients in comparison to patients with pelvic pain of other origins. J Psychosom Res 1993;37(2):111-116. [doi: 10.1016/0022-3999(93)90077-S] [Medline: 8463987]

11. Lorencatto C, Vieira M, Pinto C. Evaluation of the frequency of depression in patients with endometriosis and pelvic pain. Evaluation of the frequency of depression in patients with endometriosis and pelvic pain Rev Assoc Med Bras 2002:48-221. [doi: 10.1590/S0104-42302002000300033]
12. Cox H, Ski C, Wood R, Sheahan M. Endometriosis, an unknown entity: the consumer's perspective. Int J of Consum Stud Jun 1 2003;27(3):200-209. [doi: 10.1046/j.1470-6431.2003.00302.x]
13. Jones G, Jenkinson C, Kennedy S. The impact of endometriosis upon quality of life: a qualitative analysis. Journal of Psychosomatic Obstetrics & Gynecology 2009 Jul 07;25(2):123-133. [doi: 10.1080/01674820400002279] [Medline: 15715035]
14. Friedl F, Riedl D, Fessler S, Wildt L, Walter M, Richter R, et al. Impact of endometriosis on quality of life, anxiety, and depression: an Austrian perspective. Arch Gynecol Obstet 2015 Dec;292(6):1393-1399. [doi: 10.1007/s00404-015-3789-8] [Medline: 26112356]
15. Fourquet J, Báez L, Figueroa M, Iriarte RI, Flores I. Quantification of the impact of endometriosis symptoms on health-related quality of life and work productivity. Fertil Steril 2011 Jul;96(1):107-112 [FREE Full text] [doi: 10.1016/j.fertnstert.2011.04.095] [Medline: 21621771]
16. Gilmour JA, Huntington A, Wilson HV. The impact of endometriosis on work and social participation. Int J Nurs Pract 2008 Dec;14(6):443-448. [doi: 10.1111/j.1440-172X.2008.00718.x] [Medline: 19126072]
17. Nnoaham KE, Hummelshoj L, Webster P, d'Hooghe T, de Cicco Nardone Fiorenzo, de Cicco Nardone Carlo, World Endometriosis Research Foundation Global Study of Women's Health consortium. Impact of endometriosis on quality of life and work productivity: a multicenter study across ten countries. Fertil Steril 2011 Aug;96(2):366-373.e8 [FREE Full text] [doi: 10.1016/j.fertnstert.2011.05.090] [Medline: 21718982]
18. Moradi M, Parker M, Sneddon A, Lopez V, Ellwood D. Impact of endometriosis on women's lives: a qualitative study. BMC Womens Health 2014;14:123 [FREE Full text] [doi: 10.1186/1472-6874-14-123] [Medline: 25280500]
19. Denny E, Mann CH. Endometriosis-associated dyspareunia: the impact on women's lives. J Fam Plann Reprod Health Care 2007 Jul;33(3):189-193 [FREE Full text] [doi: 10.1783/147118907781004831] [Medline: 17609078]
20. Whitney ML. Importance of lay organizations for coping with endometriosis. J Reprod Med 1998 Mar;43(3 Suppl):331-334. [Medline: 9564669]
21. Martins MV, Peterson BD, Almeida VM, Costa ME. Direct and indirect effects of perceived social support on women's infertility-related stress. Hum Reprod 2011;26(8):2113-2121. [doi: 10.1093/humrep/der157]
22. Dukes HK, Holahan C. The relation of social support and coping to positive adaptation to breast cancer. Psychol Health Jan 1 2003;18(1):15-29. [doi: 10.1080/0887044031000080656]
23. Markovic M, Manderson L, Warren N. Endurance and contest: women's narratives of endometriosis. Health (London) 2008 Jul;12(3):349-367. [doi: 10.1177/1363459308090053] [Medline: 18579632]
24. Cox H, Henderson L, Andersen N, Cagliarini G, Ski C. Focus group study of endometriosis: struggle, loss and the medical merry-go-round. Int J Nurs Pract 2003 Feb;9(1):2-9. [doi: 10.1046/j.1440-172X.2003.00396.x] [Medline: 12588614]
25. Matthews D. Internet Access - Households and Individuals. 2013. URL: http://www.ons.gov.uk/peoplepopulationandcommunity/householdcharacteristics/homeinternetandsocialmediausage/bulletins/internetaccesshouseholdsandindividuals/2013-08-08 [accessed 2016-04-25] [WebCite Cache ID 6h2IMKUGQ]
26. Fox S, Duggan M. Health online. URL: http://www.pewinternet.org/2013/01/15/health-online-2013/ [accessed 2016-04-25] [WebCite Cache ID 6h2Keuu2N]
27. Potts H. Online support groups: an overlooked resource for patients. Health Info Internet Apr 1 2005;44(1):6-8.
28. Eysenbach G. The Impact of the Internet on Cancer Outcomes. CA: A Cancer Journal for Clinicians 2003 Nov 01;53(6):356-371 [FREE Full text] [doi: 10.3322/canjclin.53.6.356] [Medline: 15224975]
29. Wright KB, Bell SB, Wright KB, Bell SB. Health-related Support Groups on the Internet: Linking Empirical Findings to Social Support and Computer-mediated Communication Theory. J Health Psychol 2003 Jan;8(1):39-54. [doi: 10.1177/1359105303008001429] [Medline: 22113899]
30. Neal DM, McKenzie PJ. Putting the pieces together: endometriosis blogs, cognitive authority, and collaborative information behavior. J Med Libr Assoc 2011 Apr;99(2):127-134 [FREE Full text] [doi: 10.3163/1536-5050.99.2.004] [Medline: 21464850]
31. Whelan E. 'No one agrees except for those of us who have it': endometriosis patients as an epistemological community. Sociol Health Illn 2007 Nov;29(7):957-982 [FREE Full text] [doi: 10.1111/j.1467-9566.2007.01024.x] [Medline: 18092978]
32. van Uden-Kraan Cornelia F, Drossaert Constance H C, Taal E, Shaw BR, Seydel ER, van de Laar Mart A F J. Empowering processes and outcomes of participation in online support groups for patients with breast cancer, arthritis, or fibromyalgia. Qual Health Res 2008 Mar;18(3):405-417. [doi: 10.1177/1049732307313429] [Medline: 18235163]
33. Vilhauer RP. Perceived benefits of online support groups for women with metastatic breast cancer. Women Health 2009;49(5):381-404. [doi: 10.1080/03630240903238719] [Medline: 19851944]
34. Buchanan H, Coulson NS. Accessing dental anxiety online support groups: an exploratory qualitative study of motives and experiences. Patient Educ Couns 2007 Jun;66(3):263-269. [doi: 10.1016/j.pec.2006.12.011] [Medline: 17320336]
35. Malik SH, Coulson NS. Computer-mediated infertility support groups: an exploratory study of online experiences. Patient Educ Couns 2008 Oct;73(1):105-113. [doi: 10.1016/j.pec.2008.05.024] [Medline: 18639409]

36. Esquivel A, Meric-Bernstam F, Bernstam EV. Accuracy and self correction of information received from an internet breast cancer list: content analysis. BMJ 2006 Apr 22;332(7547):939-942 [FREE Full text] [doi: 10.1136/bmj.38753.524201.7C] [Medline: 16513686]
37. Gibson J. The ecological approach to visual perception. New York: East Sussex; Psychology Press; 2015.
38. Norman D. The psychology of everyday things. In: Basic books. New York: Basic Books; 1988.
39. Merolli M, Gray K, Martin-Sanchez F. Therapeutic affordances of social media: emergent themes from a global online survey of people with chronic pain. J Med Internet Res 2014;16(12):e284 [FREE Full text] [doi: 10.2196/jmir.3494] [Medline: 25533453]
40. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psych Jan 1 2006;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
41. Day D, Lloyd M. Affordances of online technologies: More than the properties of the technology. Austral Educ Comp 2007;22(2):17-21.
42. Zhang P. Technical opinion Motivational affordances: reasons for ICT design use. Commun ACM Nov 1 2008;51(11):145-147. [doi: 10.1145/1400214.1400244]
43. Zhao Y, Liu J, Tang J, Zhu Q. Conceptualizing perceived affordances in social media interaction design. In: AP. 2013 Mar 01 Presented at: Aslib Proceedings; 2013 Mar 1; China p. 289-303. [doi: 10.1108/00012531311330656]
44. Denny E. Women's experience of endometriosis. J Adv Nurs 2004 Jun;46(6):641-648. [doi: 10.1111/j.1365-2648.2004.03055.x] [Medline: 15154905]
45. Cox H, Henderson L, Wood R, Cagliarini G. Learning to take charge: women's experiences of living with endometriosis. Complement Ther Nurs Midwifery 2003 May;9(2):62-68. [doi: 10.1016/S1353-6117(02)00138-5] [Medline: 12697156]
46. Merolli M, Gray K, Martin-Sanchez F, Lopez-Campos G. Patient-reported outcomes and therapeutic affordances of social media: findings from a global online survey of people with chronic pain. J Med Internet Res 2015;17(1):e20 [FREE Full text] [doi: 10.2196/jmir.3915] [Medline: 25616273]

Abbreviations

PRO: patient-reported outcome

Edited by G Eysenbach; submitted 22.01.16; peer-reviewed by T Turnbull, K Hanna; comments to author 10.02.16; revised version received 22.02.16; accepted 22.02.16; published 09.05.16.

Please cite as:
Shoebotham A, Coulson NS
Therapeutic Affordances of Online Support Group Use in Women With Endometriosis
J Med Internet Res 2016;18(5):e109
URL: http://www.jmir.org/2016/5/e109/
doi:10.2196/jmir.5548
PMID:27160641

Original Paper

Web-based Discussion Forums on Pregnancy Complaints and Maternal Health Literacy in Norway: A Qualitative Study

Eva Haukeland Fredriksen[1*], PhD; Janet Harris[2*], PhD; Karen Marie Moland[3,4*], PhD

[1]Department of Occupational Therapy, Physiotherapy and Radiography, Faculty of Health and Social Sciences, Bergen University College, Bergen, Norway

[2]School of Health & Related Research, University of Sheffield, Sheffield, United Kingdom

[3]Centre for International Health, Faculty of Medicine and Dentistry, University of Bergen, Bergen, Norway

[4]Department of Nursing, Faculty of Health and Social Sciences, Bergen University College, Bergen, Norway

[*]all authors contributed equally

Corresponding Author:
Eva Haukeland Fredriksen, PhD
Department of Occupational Therapy, Physiotherapy and Radiography
Faculty of Health and Social Sciences
Bergen University College
Inndalsveien 28, Bergen
Postbox 7030, 5020 Bergen
Norway
Phone: 47 55585607
Email: eva.haukeland.fredriksen@hib.no

Abstract

Background: The Internet is one of the fastest growing information sources for pregnant women and seems to be used across social and economic strata. However, we still lack knowledge on how interaction in Web-based discussion forums influence maternal health literacy, in terms of how pregnant women access, appraise, and apply information to promote and maintain good health.

Objective: The aim of this study was to understand how Web-based discussion forums influence maternal health literacy; hence, we explored the role of interactions in Web-based discussion forums among women who experienced health problems during pregnancy. More specifically, we explored why media-literate women experiencing the medically unexplained condition, pelvic girdle pain (PGP), during pregnancy participated in Web-based discussion forums and how they appraised and applied the information and advice that they gained from the Web-based interaction with other women.

Methods: Women were invited to participate in the study via postings on 3 different open websites for pregnant women and mothers. The sample included 11 Norwegian women who participated in open Web-based discussion forums when experiencing PGP in pregnancy. The data were collected using synchronous qualitative email interviews and were analyzed using thematic analysis.

Results: In our study sample, interaction in Web-based discussion forums influenced maternal health literacy in terms of increased health-related knowledge and competencies, increased awareness of health promotion and health protection, and increased system navigation. The women appraised and selectively applied information and advice that resonated with their own experiences. For many, the information provided online by other women in the same situation was valued more highly than advice from health professionals. Women reported that they used their knowledge and competency in encounters with health professionals but hesitated to disclose the origin of their knowledge. Those with a high level of education in medicine-related fields raised a concern about the Internet as a source of horror stories and erroneous information and were actively engaged in trying to minimize potential negative effects, by providing biomedical information.

Conclusions: The popularity of Web-based discussion forums among pregnant women suggests that this group needs additional sources of information and support to complement traditional consultations with the health professionals. The professionals need to recognize that pregnant women access Web-based discussion forums for support and information to increase their ability to take better health decisions for themselves. This is a potential resource that health professionals may find useful in consultations with pregnant women.

(J Med Internet Res 2016;18(5):e113) doi:10.2196/jmir.5270

KEYWORDS

qualitative research; Internet; pregnancy; health literacy; web-based discussion forums; pelvic girdle pain

Introduction

The Internet is one of the fastest growing sources of information on pregnancy-related health topics and plays a significant role in health information seeking and decision-making, as well as for social networking and support [1,2]. People go online for different purposes. Some look for information or look for other people's stories. Others present their own stories or communicate with people in similar situations to access support and make sense of information [3-6]. People also provide online health support outside the professionals' domain [6-8], and electronic peer-to-peer support groups provide people with unprecedented opportunities to share information and become experts in their condition [4,9]. Pregnant women access the Internet to obtain information [10-12] and gain more control over decisions affecting pregnancy [10]. A study conducted in 2010 reported that the majority participated in Web-based pregnancy discussion groups more than 10 times during pregnancy [2]. A follow-up study from 2011 found that, for many women, Web-based pregnancy discussion groups assisted their decision-making [10].

In Norway, almost everyone has access to the Internet, and most people have access to more than one information and communication technology platform. According to 2015 statistics, 96% of the population in the 16-79 year age group were Internet users, and 90% of the women in the 25-34 year age group used the Internet to seek for health-related information [13]. Web-based discussions forums are popular health information sources for pregnant women; however, we lack knowledge on how they use such discussion forums to increase their health literacy.

“Health literacy” (HL) is a term introduced in the 1970s, but there is still not one universally accepted definition for this term [14]. In 1988, World Health Organization defined HL as “The cognitive and social skills which determine the motivation and ability of individuals to gain access to, understand, and use the information in ways which promote and maintain good health” [14]. Interest in Internet HL (eHealth literacy) has focused on being able to access relevant information from electronic sources and being able to make sense of it. Norman and Skinner define eHealth literacy as the ability to seek, find, understand, and appraise health information from electronic sources and apply the knowledge gained to address or solve a health problem [15]. Their definition rests on Nutbeam’s model, which describes 3 levels of HL, progressing from basic skills in reading and writing (functional HL), to the ability to derive meaning from different forms of communication and apply new information to changing situations (interactive HL), and to the ability to achieve policy and organizational changes (critical HL) [14,16-19]. Nutbeam and Renkert introduced the concept of “maternal health literacy” [20]. This concept is about interactive HL, describing the “cognitive and social skills which determine the motivation and ability of women to gain access to, understand, and use information in ways that promote and maintain their health and that of their children” [20]. In this study, we used a combination of the definitions of interactive HL and eHealth literacy to explore how interaction in Web-based discussion forums influences pregnant women's motivation and ability to access, appraise, and apply information to promote and maintain their health.

We have used the medically unexplained condition, pelvic girdle pain (PGP), as a tracer condition because an increasing number of pregnant women, especially in the Scandinavian countries, are diagnosed with PGP [21]. PGP is about pain in the pelvic girdle and lumbar regions [22] and is a common cause of sick leave among Norwegian pregnant women [23]. Women who experience PGP commonly experience diminished endurance capacity for everyday life activities, such as standing, walking, and sitting [22]. We explored why media-literate Norwegian women who experienced PGP during pregnancy participated in Web-based discussion forums and how they appraised and applied the information and advice that they gained from Web-based interactions with other women.

Methods

Synchronous Qualitative Email Interviews

We decided to collect data by using email interviews because this method of data collection enabled women from a wide geographical area to participate in interviews without making any special arrangements. Email interviews also enabled us to recruit women who were juggling work with pregnancy or motherhood and who might view online interviews as less of a burden than face-to-face interviews. We also wanted answers in a here-and-now dialogue between the interviewees and the researcher and decided to use the method of synchronous email interviews. This interview method is described in detail in the Data collection and Analysis section. Next, we developed a thematic semistructured interview guide with 5 main topics: (1) motivation for participation in Internet discussions on PGP, (2) learning outcome, (3) support, (4) encounters with the health professionals, and (5) Internet discussions and health behavior. The interview guide was used as a flexible tool during the interview process.

Sampling

Invitations to participate in the study were posted on 3 different open websites for pregnant women and mothers. We sought computer-literate women who could read, write, and understand Norwegian, who experienced PGP during pregnancy, and who had participated in Web-based discussion forums during pregnancy. However, we experienced that sampling via the Internet was more challenging than expected. It took us more than 5 months to set up a sample of 11 women who fulfilled the inclusion criteria for the study. Twenty-three women expressed interest, and 11 of them participated in the study. Among the nonparticipants, 7 did not respond to follow-up

mails, and 5 were excluded because they did not meet the inclusion criteria.

The first author received permission from the people in charge of the open website www.barnimagen.com to advertise for participants on the discussion forum on PGP, and the invitation was posted in October 2013. During the next 2 months, 10 women made contact and expressed their interest. We emailed invitations with detailed project information and a consent form, but 3 women never responded to the initial or follow-up emails. Seven women responded positively and participated in email interviews during autumn 2013. Next, we decided to post the invitation on an additional Internet site to increase the number of participants. The first author received permission from the midwife in charge of the site www.altformamma.com to advertise for participants. The invitation was posted in December 2013, and it was reposted several times over the next months to generate more informants. During this period, 13 women contacted us. Among those, 5 did not fulfill the inclusion criteria, and 4 did not respond to follow-up emails. Four women were included in the study, with interviews taking place between January and March 2014. To increase the number of participants, we also advertised for participants on the website www.babyverden.com in February 2014. However, there were no responses to our advertisement as it was difficult to see among the large number of threads and discussions on this website. After the final interviews in March, we decided to stop further sampling because although the sample was small, we had quite a rich and nuanced dataset, representing women from different social and professional backgrounds, geographical areas, age groups, and pregnancy status.

Participants

All the women had participated in Web-based discussions on PGP, and all had experienced PGP during their pregnancies. Most of them were on maternity leave, or were sick-listed due to PGP, and many experienced persisting symptoms of PGP after delivery. We have given nicknames to the participants so that the readers can discriminate between them while reading the Result section.

Table 1. Participants.

Participant	Age, years	Profession	Pregnancy status	Nonpregnant women: time since delivery	PGP[a] experience
Anne	28	Secretary	First pregnancy		PGP in pregnancy, sick-listed
Betty	30	Engineer	Second pregnancy		PGP in both pregnancies, sick-listed
Cecilia	29	Nurse	Not pregnant, 2 children	16 weeks	PGP in both pregnancies, persisting symptoms
Dea	24	Bank employee	Third pregnancy		PGP in her last 2 pregnancies
Eva	36	Veterinary	Second pregnancy		PGP in both pregnancies
Frida	39	Assistant nurse	Not pregnant, 2 children	4 years	PGP in both pregnancies, persisting symptoms
Gill	39	Medical doctor	Not pregnant, 1 child	18 months	PGP in pregnancy, persisting symptoms
Hannah	22	Health care assistant	Not pregnant, 1 child	3 months	PGP in pregnancy, persisting symptoms
Inga	38	Farmer	Not pregnant, 1 child	18 months	PGP in pregnancy, persisting symptoms, sick-listed
Jenny	36	Cleaner	Not pregnant, 3 children	6 months	PGP in all pregnancies, persisting symptoms
Karen	22	Kiosk worker	Not pregnant, 1 child	6 months	PGP in pregnancy

[a]PGP: pelvic girdle pain.

Data Collection and Analysis

All email interviews were carried out in Norwegian by the first author. The interview situation lacked the body language, gestures, smiles, eye contact, and small talk that occur in face-to-face interviews; hence, it demanded a different approach to establish trust and dialogue. The researcher sent follow-up responses with encouraging comments to the participants during the interviews to establish trust and facilitate communication. However, extra care was taken about how the conversation was phrased throughout the interviews, as the researcher was aware that the research participants might use the email correspondence for their own purposes.

One topic was addressed at a time. After receiving the answers on 1 topic, follow-up questions were sent to ask for more information or to prevent misunderstandings, before turning to the next topic. Sometimes, short summaries were written, asking the woman to confirm her interpretation of their answers. This worked well. Many women contributed with detailed answers, and the follow-up questions generated additional information on each topic. Most interviews took 1 to 2 hours, and there were no dropouts. One interview was interrupted by grocery shopping, and another was delayed owing to sick children. These incidents

were solved by taking pauses or by completing the interview the next day.

The first author transferred the email correspondence to text files after the interviews, and then she deleted the mails. The Norwegian authors read the interview documents in full text, before the first author performed the initial coding of the data, using Nvivo10. In the initial part of the analysis process, the first author used the concepts of interactive HL [15-17,20] to guide the analysis and identify themes, looking for topics related to motivation and ability to seek, find, understand, appraise, and apply Web-based discussion forum information to changing situations, in ways that promote and maintain health during pregnancy. Thereafter, the team worked together, performed a thematic analysis, based on the initial organization of the data, and reorganized the data under new themes and key themes.

Ethics Approval

Written informed consent was obtained from all the participants. Ethical approval was given by the Regional Committee for Research Ethics (reference number 2012/2225/REK Vest). We have carefully sheltered the anonymity of all the participants in the study, and the data are stored in a password-based research server, accessible only to the research team.

Results

Key Themes

After the analysis process described earlier, we organized the findings into 3 key themes: (1) seeking experience-based health information online; (2) understanding and appraising experience-based online health information; and (3) taking control over one' health. The quotes that illustrate the findings were translated to English by the Norwegian authors.

Seeking Experience-Based Health Information Online

This theme describes why the participants looked for experience-based health information online and how they judged the relevance of this information. The theme resonates with central dimensions of interactional HL, such as seeking, finding, and appraising relevance of information from different forms of communication.

Most women went online to look for people in the same situation as themselves. Some were primarily looking for information, whereas others were looking for emotional support, detailed advice about special topics, or alternative courses of action for themselves. The expectations and motives to participate in the Internet discussions differed across the sample according to previous experience with PGP. Those without previous experience with PGP, such as Hannah and Karen, sought information and emotional support. Hannah attended the discussion forum because she was worried and needed advice: "*It is good to read that other women are in the same situation (…). If I wonder about something, I always get answers.*" Karen looked for information and practical ideas from women in the same situation:

> *"I got advice from my mom and my mother in law, but it is a long time since they had babies and they may have forgotten a lot. I feel much better discussing my problems with women who are in the same situation.*"

However, pregnant women who had former experience of PGP also sought support and advice from others. Betty wrote: "*I want to meet women who have similar experiences so that I can learn how they managed their pain and what kind of adjustments they made. For instance, what was their threshold for requesting sick leave?*"

Dea participated to diagnose her own symptoms, and to "*receive advice from others who are, or have been in, a similar situation. Above all, I participate because it is best to get advice from people in the same situation as myself.*" Others, as Inga, were primarily looking for emotional support: "*I participate to feel less lonely and to talk with others in the same situation as myself.*"

Others wanted to give support and share knowledge with others, to help them solve their health problems. Many women who gave medical information and advice to others had no professional background from medicine-related fields. The engineer, Betty, for example, said: "*The reason why I have participated in the Internet discussions is that I want to help other women so that they do not have to go through the same distressing experience as I did.*"

Women such as Eva, with high education in medicine-related fields, justified their participation differently: "*I answer if I can contribute with good information. I know much about the medical background of PGP and I have also experienced PGP myself.*"

The women found what they were looking for, appraising the information and support as relevant because it was experience-based and updated and originated from women in the same situation as themselves.

Understanding and Appraising Experience-Based Online Health Information

This theme describes how interaction with others influenced the participants' knowledge on how to deal with their health problems and how they appraised the validity of the health information that they derived from Web-based interactions. Thus, the theme resonates with other dimensions of interactional HL, such as understanding and appraising validity of information to address or solve a health problem.

The discussion forums seemed to contribute to increased knowledge and higher capability to act on health problems. The women informed each other about symptoms, incidence, and prognosis. Dea said that the Web-based discussions were vital for her knowledge about PGP: "*I think that as much as 80% of my first knowledge about PGP was via other women online.*"

However, this knowledge also made them more aware of potential health risks and of risk reduction. Women reported that the Internet forums had informed them that pushing their physical activity limits during pregnancy might increase the risk of developing persisting health problems after delivery. Such stories fostered an awareness of risks. Women posted directions on where to seek help and treatment and how to request health care or sick leave. However, the advice-seeking

women reported that they compared other women's experiences to their own experiences and symptoms before they decided whether to implement the advice.

For many, the advice from others in the same situation fostered an increased understanding of how to maintain health and well-being. Betty had learnt "*what to do to avoid massive health problems in this pregnancy,*" and Karen thought that her new knowledge might help her in her next pregnancy, "*Then I know what to do to avoid massive health problems such as those that I have experienced during this pregnancy. I will be more prepared, so to say.*" Moreover, they had learnt the importance of requesting professional help to prevent worsening of the symptoms. Betty commented that she had learnt what others normally received help for and what she might request from the doctor. Anne described how the Internet forum had given her information on where to find help to cope with the situation and how to describe her symptoms to get adequate health care: "*I have learnt what to tell to receive the help that I need.*"

Women also shared information about social rights, such as access to treatment, physical remedies, job adjustments, or sick leave. This information was highly valued. Some commented that the doctors often lacked an overview of social rights, and that women need to know their rights to receive adequate help. Inga, among others, blamed herself for the persisting health problems: "*If I had known earlier what I know now, I would have demanded more support from the health professionals. Then I could have coped better with the problems.*"

A few women, such as Eva and Gill, who both had high education in medicine-related fields, were concerned about the validity of experience-based knowledge. Eva said that she contributed to the discussions with what she termed "correct information": "*Sometimes people write horror-stories, and contribute with incorrect information. Then I think it is important to add correct information to the discussions.*" Gill commented that she did not find the information on the discussion forums trustworthy unless it was linked to reliable sources. Other women were less concerned about validity:

> *"Some answers might have poor quality. But I think that there is valuable advice online, especially advice from others in the same situation who have new experiences. I think that this advice is often better than advice from health professionals, who lack this bodily experience themselves.*" *(Dea)*

Most women developed increased understanding of how to maintain health and well-being in pregnancy, and they appraised the validity of experience-based advice from peers.

Taking Control Over Own Health

This theme describes how the participants took control over their own health by applying information to promote and maintain their health and illustrates another central dimension of maternal HL.

Many participants described how they navigated in the health system, how they negotiated with the health professionals to get their needs met, and how they applied the advice from the Internet to cope with their everyday life. Some reported how they had contacted doctors and requested physiotherapy, chiropractic services, or acupuncture treatment. Anne, for example, wrote: "*Participation in these discussions has convinced me that I have PGP. That is the reason why I took the next step, and contacted a doctor.*" Jenny negotiated with her doctor to receive treatment for PGP while pregnant, arguing that she was eligible to the same social rights as other women in her situation. She also said that she had learnt what to request if her health problems worsened. Other women shared this experience of being empowered. They commented that support and advice from other women had made them more determined to request help, adjustment, and treatment. Frida said that when she had requested physiotherapy treatment, she did not give in until the doctor responded to her requests. Eva thought that because of her background, she had different experiences: "*My impression is that they listen differently and better to me because I have medical education, which enables me to argue for my needs.*" Nevertheless, she also valued information on treatment options and had followed advice to seek physiotherapy treatment. The medical doctor, Gill, found it interesting to participate in the Internet discussions, but she was clear that this did not influence her health behavior.

However, some women did not argue for their needs at all. They said that they hid their knowledge and its origins in the encounter with the health care providers because they knew that the professionals did not value experience-based knowledge from Internet forums. Dea told that she never shared her thoughts and knowledge with the health care providers because she had experienced that the professionals did not listen to her: "*I am only saying that I know nothing.*" Karen also kept her knowledge to herself: "*I have never received positive feedback from my doctor when I have told what I have read on the Internet (…). He said that most online information was wrong.*" Dea and Karen took control over their health by using a strategy of covert resistance vis-a-vis the health professionals, using advice from others that resonated with their own experiences. Hannah also reported how Internet discussions had enabled her to get better health control. She reported that she had learnt that she should relax and let other people do things for her:

> *"I have good health control now, as a result of good advice from other women on the Internet. (…) I have used good exercises and pelvic belt, relaxed when I had pain, and accepted that it is ok to stay in bed if needed.*"

The data indicate selective application of knowledge, as the women appraised and applied information and advice that resonated with their own experiences.

Discussion

Principal Findings

In our study sample, interaction in Web-based discussion forums influenced maternal HL, in terms of increased health-related knowledge and competencies, increased awareness of health promotion and health protection, and better capability to navigate health systems. Women critically appraised and selectively applied information and advice that resonated with their own experiences. For many, the information provided

online from other women in the same situation was valued more highly than advice from health professionals. Women reported that they used their knowledge and competency in the encounter with health professionals but hesitated to disclose the origin of their knowledge. Those with a high level of education in medicine-related fields raised a concern about the discussion forums as a source of horror stories and erroneous information and were actively engaged in trying to minimize potential negative effects, by providing biomedical information. Women's concerns about the quality of online pregnancy-related health information have been reported elsewhere [10]. It has also been reported that personal stories from other people in the same situation are not necessarily reliable and can represent both a negative and a positive source of support [24]. However, in our study, concerns about the quality and trustworthiness of the information were generally overshadowed by the positive evaluation of experience-based knowledge and support from others in the same situation.

Our findings on Web-based discussion forums as a source of information and support for pregnant women resonate with findings of other studies in the field [6,10,24]. A study of the role of the Internet in problematic pregnancies reported that women used the Internet as a complementary source for pregnancy information and also to learn from others in similar situations [24]. Several studies have reported that pregnant women turned to technology to fill their knowledge gap [2,10-12,25] and that more informed patients can change the interaction with the health professionals [4,5]. However, previous research has reported diverging findings on whether pregnant women revealed their Internet-based knowledge and discussed this information with the health professionals [2,10-12]. A Web-based survey from 2010, with participants from 24 countries, found that 70% of the participants referred to Internet information in discussion with at least 1 health care provider and that almost 90% of those reported that the information was welcomed and openly discussed [2]. Other studies, from China and Sweden, reported that the majority had not discussed information from the Internet with their health care providers [11,12]. Diverging findings on communication with the health professionals may refer to different cultural norms in terms of the types of knowledge people bring into the medical encounter. However, when discussing communication with the health professionals, the studies referred previously [2,10-12] did not distinguish between different information sources on the Internet. Hence, we do not know whether the women only discussed information that they had retrieved from official health information sites or whether they also discussed information that they had retrieved from informal Web-based discussion forums. Our finding that women actively hide discussion forums as their sources of information was not discussed in the aforementioned studies.

As far as we know, previous Internet research has not explored the influence of informal discussion forums on maternal HL. Women have traditionally shared private concerns and received informal support and advice from other women during pregnancy. This sharing and caring in an informal private space has partly been replaced by informal interaction on the Internet. Interaction with others who share the same illness experiences may have greater weight than advice from health professionals, who are unlikely to have these experiences [6]. Women also tend to listen to other women in the same situation when their embodied experiences diverge from the experts' biomedical knowledge [26,27]. Many participants in our study valued the Web-based information from other women higher than advice from health professionals when the information resonated with their own experiences. The women found the information and support enabling, in terms of controlling their health condition. It has been reported elsewhere that Web-based support can give people a sense of empowerment [6,10,24], and there is increasing evidence that hearing other people's stories can affect health behavior [6]. In general, support from others may have positive health effects because it may increase motivation for self-care and encourage help-seeking behavior and may thus prevent minor illnesses from developing into more serious conditions [28]. However, peer support may also encourage healthy people to request medical assistance for normal life events and minor health problems and may thus contribute to increased medicalization [28]. More seriously, hearing about other people's experiences may reinforce unhealthy behavior [2,6], influencing others to make inappropriate decisions about the management of their symptoms. However, our findings indicate that pregnant women critically appraised and selectively applied information and advice from the Internet and used the advice that resonated with their own experiences.

Strengths, Challenges, and Limitations

Using email interviews produced a sample that included women from a wide geographical area and different social backgrounds. Many participants commented that communication via email enabled them to participate from their home without making any special arrangements. Other Internet studies on pregnancy and birth have also reported that the participants valued email interviews [29].

The recruitment process was challenging. We tried to increase the number of participants by advertising for participants on several websites. However, this generated answers from women who did not fulfill the inclusion criteria. It was also challenging to create commitment online. Many withdrew from further interaction when they realized that they were expected to follow formal research procedures, including signing a written consent form. Although participants were self-selected, they had in-depth experience of the phenomenon and came from diverse backgrounds. But, the sample was weighted toward women with persisting health problems after pregnancy. We do not know whether currently pregnant women would have different views or whether different views and experiences vary systematically among women from different social and educational backgrounds.

Conclusions and Implications

In our study sample, interaction in Web-based discussion forums influenced maternal HL. The popularity of Web-based discussion forums among pregnant women suggests that this group needs additional sources of information and support to complement traditional consultations with the health professionals. Health care providers should acknowledge Web-based discussion forums as a source of health information

that may influence pregnant women's health behavior. They also need to recognize that pregnant women access Web-based discussion forums for support and information to increase their ability to take better health decisions for themselves. This is a potential resource that health professionals may find useful in consultations with pregnant women.

Acknowledgments

This study was funded by Bergen University College, Bergen, Norway. The authors also acknowledge the help of all the 11 participants, and the persons in charge of the open websites.

Conflicts of Interest

None declared.

References

1. Lagan B, Sinclair M, Kernohan W. Pregnant women`s use of the internet: A review of published and unpublished evidence. Evidence Based Midwifery 2006;4(1):17-23.
2. Lagan BM, Sinclair M, Kernohan WG. Internet use in pregnancy informs women's decision making: a web-based survey. Birth 2010 Jun;37(2):106-115. [doi: 10.1111/j.1523-536X.2010.00390.x] [Medline: 20557533]
3. Hardey M. Doctor in the house: the Internet as a source of lay health knowledge and the challenge to expertise. Sociol Health & Illness 1999 Nov;21(6):820-835. [doi: 10.1111/1467-9566.00185]
4. Broom A. Virtually he@lthy: the impact of internet use on disease experience and the doctor-patient relationship. Qual Health Res 2005 Mar;15(3):325-345. [doi: 10.1177/1049732304272916] [Medline: 15761103]
5. Sommerhalder K, Abraham A, Zufferey MC, Barth J, Abel T. Internet information and medical consultations: experiences from patients' and physicians' perspectives. Patient Educ Couns 2009 Nov;77(2):266-271. [doi: 10.1016/j.pec.2009.03.028] [Medline: 19411157]
6. Ziebland S, Wyke S. Health and illness in a connected world: how might sharing experiences on the internet affect people's health? Milbank Q 2012 Jun;90(2):219-249 [FREE Full text] [doi: 10.1111/j.1468-0009.2012.00662.x] [Medline: 22709387]
7. Hardey M. 'The Story of My Illness': Personal Accounts of Illness on the Internet. Health: 2002 Jan 01;6(1):31-46. [doi: 10.1177/136345930200600103]
8. Pitts V. Illness and Internet empowerment: writing and reading breast cancer in cyberspace. Health (London) 2004 Jan;8(1):33-59. [doi: 10.1177/1363459304038794] [Medline: 15018717]
9. Barker KK. Electronic support groups, patient-consumers, and medicalization: the case of contested illness. J Health Soc Behav 2008 Mar;49(1):20-36. [Medline: 18418983]
10. Lagan BM, Sinclair M, Kernohan WG. What is the impact of the Internet on decision-making in pregnancy? A global study. Birth 2011 Dec;38(4):336-345. [doi: 10.1111/j.1523-536X.2011.00488.x] [Medline: 22112334]
11. Gao L, Larsson M, Luo S. Internet use by Chinese women seeking pregnancy-related information. Midwifery 2013 Jul;29(7):730-735. [doi: 10.1016/j.midw.2012.07.003] [Medline: 22958935]
12. Larsson M. A descriptive study of the use of the Internet by women seeking pregnancy-related information. Midwifery 2009 Feb;25(1):14-20. [doi: 10.1016/j.midw.2007.01.010] [Medline: 17408822]
13. Statistics Norway. Statistics Norway. Oslo: SSB; 2015. ICT usage in households URL: http://www.ssb.no/statistikkbanken [accessed 2016-04-25] [WebCite Cache ID 6h200ribr]
14. Sørensen K, Van den Broucke Stephan, Fullam J, Doyle G, Pelikan J, Slonska Z, (HLS-EU) Consortium Health Literacy Project European. Health literacy and public health: a systematic review and integration of definitions and models. BMC Public Health 2012;12:80 [FREE Full text] [doi: 10.1186/1471-2458-12-80] [Medline: 22276600]
15. Norman CD, Skinner HA. eHealth Literacy: Essential Skills for Consumer Health in a Networked World. J Med Internet Res 2006;8(2):e9 [FREE Full text] [doi: 10.2196/jmir.8.2.e9] [Medline: 16867972]
16. Nutbeam D. Health literacy as a public health goal: a challenge for contemporary health education and communication strategies in the 21st century. Health Promotion International 2000;15:259-267. [doi: 10.1093/heapro/15.3.259]
17. Nutbeam D. The evolving concept of health literacy. Soc Sci Med 2008 Dec;67(12):2072-2078. [doi: 10.1016/j.socscimed.2008.09.050] [Medline: 18952344]
18. Gillis D. University of Nottingham. Nottingham: University of Nottingham; 2009. Exploring dimensions of health literacy: A case study of interventions to promote and support breast feeding URL: http://eprints.nottingham.ac.uk/11199/ [accessed 2016-04-30] [WebCite Cache ID 6h9Myiulb]
19. Chinn D. Critical health literacy: a review and critical analysis. Soc Sci Med 2011 Jul;73(1):60-67. [doi: 10.1016/j.socscimed.2011.04.004] [Medline: 21640456]
20. Renkert S. NutbeamD. Opportunities to improve maternal health literacy through antenatal education: an exploratory study. Health promotion International (4) 2001;16:381-388. [doi: 10.1093/heapro/16.4.381]

21. Bastiaanssen JM, de Bie Rob A, Bastiaenen Caroline H G, Essed Gerard G M, van den Brandt Piet A. A historical perspective on pregnancy-related low back and/or pelvic girdle pain. Eur J Obstet Gynecol Reprod Biol 2005 May 1;120(1):3-14. [doi: 10.1016/j.ejogrb.2004.11.021] [Medline: 15866079]
22. Vleeming A, Albert H, Østgaard H, Stuge B, Bengt S. Contract no.: WG4. Brussels: European Commission, Research Directorate- General, Departement of Policy, Coordination and Strategy; 2004. European Guidelines on the diagnosis and treatment for pelvic girdle pain URL: http://almacen-gpc.dynalias.org/publico/Dolor%20pelvico%20European%20Guidelines.pdf [accessed 2016-04-25] [WebCite Cache ID 6h21dnDVI]
23. Myklebø S. Sykefravaer og svangerskap [Sick leave and pregnancy]. Arbeid og velferd 2007;3:21-27.
24. Lowe P, Powell J, Griffiths F, Thorogood M, Locock L. Making it all normal: the role of the internet in problematic pregnancy. Qual Health Res 2009 Oct;19(10):1476-1484. [doi: 10.1177/1049732309348368] [Medline: 19805809]
25. Kraschnewski J, Chuang C, Poole E, Peyton T, Blubaugh I, Pauli J, et al. Paging "Dr. Google": does technology fill the gap created by the prenatal care visit structure? Qualitative focus group study with pregnant women. J Med Internet Res 2014;16(6):e147 [FREE Full text] [doi: 10.2196/jmir.3385] [Medline: 24892583]
26. Abel, EK, Browner, CH. Selective compliance with biomedical authority and the uses of experimental knowledge. In: Lock M, Kaufert P, editors. Pragmatic women and body politics. New York: Cambridge University Press; 1998.
27. Kaufert, P.. Women, resistance and the breast cancer movement. In: Lock M, Kaufert P, editors. Pragmatic women and body politics. New York: Cambridge University Press; 1998.
28. Dennis C. Peer support within a health care context: a concept analysis. Int J Nurs Stud 2003 Mar;40(3):321-332. [Medline: 12605954]
29. Beck CT. Benefits of participating in internet interviews: women helping women. Qual Health Res 2005 Mar;15(3):411-422. [doi: 10.1177/1049732304270837] [Medline: 15761108]

Abbreviations

PGP: pelvic girdle pain
HL: health literacy

Edited by A Moorhead; submitted 27.10.15; peer-reviewed by P Lowe, K Hameen-Anttila, S Meaney; comments to author 23.12.15; revised version received 02.02.16; accepted 21.02.16; published 26.05.16.

Please cite as:
Fredriksen EH, Harris J, Moland KM
Web-based Discussion Forums on Pregnancy Complaints and Maternal Health Literacy in Norway: A Qualitative Study
J Med Internet Res 2016;18(5):e113
URL: http://www.jmir.org/2016/5/e113/
doi:10.2196/jmir.5270
PMID:27230094

Original Paper

What it Takes to Successfully Implement Technology for Aging in Place: Focus Groups With Stakeholders

Sebastiaan Theodorus Michaël Peek[1,2], BICT, MSc; Eveline JM Wouters[1], MD, MSc, PhD; Katrien G Luijkx[2], MSc, PhD; Hubertus JM Vrijhoef[2,3,4], MSc, PhD

[1]Institute of Allied Health Professions, Chair of Health Innovations and Technology, Fontys University of Applied Sciences, Eindhoven, Netherlands

[2]School of Social and Behavioral Sciences, Department of Tranzo, Tilburg University, Tilburg, Netherlands

[3]Saw Swee Hock School of Public Health, National University of Singapore, Singapore, Singapore

[4]National University Health System, Singapore, Singapore

Corresponding Author:
Sebastiaan Theodorus Michaël Peek, BICT, MSc
Institute of Allied Health Professions
Chair of Health Innovations and Technology
Fontys University of Applied Sciences
Dominee Theodor Fliednerstraat 2
Eindhoven, 5631 BN
Netherlands
Phone: 31 619412612
Fax: 31 885075599
Email: research@sebastiaanpeek.nl

Abstract

Background: There is a growing interest in empowering older adults to age in place by deploying various types of technology (ie, eHealth, ambient assisted living technology, smart home technology, and gerontechnology). However, initiatives aimed at implementing these technologies are complicated by the fact that multiple stakeholder groups are involved. Goals and motives of stakeholders may not always be transparent or aligned, yet research on convergent and divergent positions of stakeholders is scarce.

Objective: To provide insight into the positions of stakeholder groups involved in the implementation of technology for aging in place by answering the following questions: What kind of technology do stakeholders see as relevant? What do stakeholders aim to achieve by implementing technology? What is needed to achieve successful implementations?

Methods: Mono-disciplinary focus groups were conducted with participants (n=29) representing five groups of stakeholders: older adults (6/29, 21%), care professionals (7/29, 24%), managers within home care or social work organizations (5/29, 17%), technology designers and suppliers (6/29, 21%), and policy makers (5/29, 17%). Transcripts were analyzed using thematic analysis.

Results: Stakeholders considered 26 different types of technologies to be relevant for enabling independent living. Only 6 out of 26 (23%) types of technology were mentioned by all stakeholder groups. Care professionals mentioned fewer different types of technology than other groups. All stakeholder groups felt that the implementation of technology for aging in place can be considered a success when (1) older adults' needs and wishes are prioritized during development and deployment of the technology, (2) the technology is accepted by older adults, (3) the technology provides benefits to older adults, and (4) favorable prerequisites for the use of technology by older adults exist. While stakeholders seemed to have identical aims, several underlying differences emerged, for example, with regard to who should pay for the technology. Additionally, each stakeholder group mentioned specific steps that need to be taken to achieve successful implementation. Collectively, stakeholders felt that they need to take the leap (ie, change attitudes, change policies, and collaborate with other organizations); bridge the gap (ie, match technology with individuals and stimulate interdisciplinary education); facilitate technology for the masses (ie, work on products and research that support large-scale rollouts and train target groups on how to use technology); and take time to reflect (ie, evaluate use and outcomes).

Conclusions: Stakeholders largely agree on the direction in which they should be heading; however, they have different perspectives with regard to the technologies that can be employed and the work that is needed to implement them. Central to

these issues seems to be the tailoring of technology or technologies to the specific needs of each community-dwelling older adult and the work that is needed by stakeholders to support this type of service delivery on a large scale.

(J Med Internet Res 2016;18(5):e98) doi:10.2196/jmir.5253

KEYWORDS

aged; independent living; technology; eHealth; health services for the elderly; project and people management; implementation management; qualitative research; focus groups

Introduction

A key challenge for most, if not all, countries is how to accommodate and care for an aging population [1]. As a response, many countries have shifted their priorities and resources toward deinstitutionalization in order to create communities that facilitate seniors to remain living in their homes for as long as possible [2]. Policies and programs that represent this paradigm shift frequently emphasize the deployment of technology as a means of supporting aging in place. Examples of technologies mentioned are sensor-based networks for activity monitoring, emergency help systems, and online tools to support older adults' self-management of chronic conditions [3,4]. These technologies are often information and communications technology (ICT) based; they are referred to as eHealth, ambient assisted living technology, smart home technology, and/or gerontechnology. Unfortunately, the implementation of these technologies is frequently unsuccessful in daily practice [5-7].

Several factors hinder the implementation of the aforementioned technologies, including low adoption levels among potential users [3,4,7,8], difficulties in building sustainable business cases [9,10], a lack of interoperability between systems of different vendors [6,9,11], and scarcity of robust scientific evidence on cost and outcomes [12-14]. All the aforementioned factors are complicated by the fact that multiple stakeholders are involved [9,15]. Typical stakeholders include older adults, care professionals, managers within home care or social work organizations, technology designers and suppliers, and policy makers. The goals and motives of these groups of stakeholders may not always be transparent or aligned [16,17]. However, empirical studies providing insight into the convergent and divergent perspectives of stakeholders involved in implementing technology that could support aging in place are few and far between. Furthermore, the few existing studies limit their focus on perceived barriers to a successful implementation [18,19] rather than forming a more complete understanding of stakeholders' positions. For example, several authors have noted that it is crucial to understand what the different stakeholders' goals are in initiatives centered around supporting aging in place with technology [20-22]. Hence, this study seeks to provide insight into the positions of stakeholder groups involved in the implementation of technology for aging in place by asking the following three questions: What kind of technology do they see as relevant for aging in place? What do they aim to achieve by implementing technology? What is needed to achieve successful implementations? A better understanding of the positions of various stakeholder groups is expected to contribute to the successful implementation of technological interventions aimed at supporting aging in place [11,20,23,24].

Methods

Sampling

This study was conducted in the Netherlands. In 2012, our research group, in collaboration with 13 partners, initiated a project aimed at finding ways to successfully deploy technologies that could support aging in place, by conducting a longitudinal field study among community-dwelling older adults. As a part of the project, five mono-disciplinary focus groups were conducted simultaneously with participants representing five groups of stakeholders within the process of implementing technology for aging in place: older adults, care professionals, managers within home care or social work organizations, technology designers and suppliers, and policy makers. These focus group sessions took place in February 2012, and convenience sampling was used by the partners of the project to recruit participants. This means that participants in the focus groups were either working for one of the partners in the project or were professional relations of partners. At the time the focus group sessions were conducted, participants representing different stakeholder groups were not engaged in implementing technology for aging in place together. Mono-disciplinary focus groups were employed because this data collection method was expected to efficiently enable productive discussions and the elicitation of a multiplicity of views by each stakeholder group [25]. Furthermore, we wanted to provide a safe environment for participants [25].

Procedure

Focus group sessions took place simultaneously in the Fontys Institute of Allied Health Professions, which is located in Eindhoven, the Netherlands. Sessions lasted 90 minutes and each session was supervised by a moderator and an assistant. Moderators had a professional background that was related to the background of the participants in their session. At the beginning of the sessions, a scenario was read out loud by the moderators. The scenario described how population aging increases the need for creative solutions to be able to continue to provide good quality care for older adults. Furthermore, the scenario explained that more and more older adults are expected to age in place, and that technological solutions are expected to play an important role in this respect. In the group discussion that followed within each session, three open-ended questions were discussed by participants. First, participants were asked what kinds of technologies they considered as "technologies that could support aging in place". This question was asked to make transparent what stakeholders perceived as technology

relevant to the context of aging in place. Second, participants were asked when they would consider the use of technology for aging in place a success. This was asked to determine what stakeholders are trying to achieve with regard to the implementation of technology for aging in place. Third, participants were asked what they need to be able to successfully implement the technology for aging in place, and what they can contribute in order to achieve successful implementations. This was done to let participants reflect on their role as stakeholders. After each question, participants were first requested to write down their answers on a form to enable them to collect their thoughts prior to engaging in the discussions. Informed consent was acquired from all participants and each session was recorded by audio and video to enable transcription. Transcripts were made anonymous and all data was only used in this study. Dutch law does not require medical or ethical reviews for focus group interviews with stakeholders other than patients. All moderators were trained according to guidelines described by Sim [25] and provided with a guide that was produced by the lead author. Each moderator was accompanied by an assistant who took notes and aided in facilitating an open dialogue between group members. Immediately after the sessions, the moderators and assistants gathered to evaluate the discussions. The moderator and assistant of the session that consisted of technology designers and suppliers stated that they had to intervene regularly because some participants were dominant in the discussion, and because participants needed to be reminded to reflect on their own role instead of focusing on the role of other stakeholders. Moderators and assistants of the other group sessions did not experience these issues, or to a far lesser extent.

Analysis

Verbatim transcripts of the sessions were analyzed using thematic analysis [26]. First, inductive codes were attached to quotations relevant to the research questions. In this process, each transcript was initially coded independently by two researchers, who subsequently had to come to an agreement and produce a single coded version of each transcript. Afterward, overarching categories of codes (ie, themes and subthemes) were formed. Additionally, the technologies that the participants deemed relevant for aging in place were classified in application domains that are part of the gerontechnology taxonomy as proposed by van Bronswijk, Bouma, and Fozard [27]. This taxonomy was selected because it is targeted toward technologies that are relevant to older adults and because it allows for the inclusion of a wide range of technologies, which is in line with the participants' responses. As a member check, a separate meeting was organized in which preliminary findings were presented. In this way, participants were provided with the opportunity to learn more about the positions of the various stakeholder groups involved in the project. Two-thirds of the participants attended the meeting and they accepted the presented findings as accurate and complete.

Results

Participants

A total of 29 participants were involved in the study and each stakeholder group was represented by 5-7 participants (see Table 1). Participants were 32-76 years old, and the average age was highest in the focus group with older adults. The managers in the study were all women. Care professionals were predominantly women, while technologists were predominantly men.

Table 1. Stakeholders and participants involved in mono-disciplinary focus groups (n=29).

Stakeholder	Description of participants	Participant characteristics	n (%)
Older adults (O)	Community-dwelling older adults (active in community voluntary work)	3 men and 3 women, aged 62-76 years	6 (21)
Care professionals (C)	Care professionals who provide home care themselves, or coordinate the provision of home care	1 man and 6 women, aged 32-55 years	7 (24)
Managers (M)	Managers within home care or social work organizations	5 women, aged 37-61 years	5 (17)
Technologists (T)	Professionals who work for companies that produce and supply technology, or for educational institutions with a focus on technology	5 men and 1 woman, aged 36-66 years	6 (21)
Policy makers or advisors to policy makers (P)	Public officers, and advisors and researchers involved in health policy	3 men and 2 women, aged 32-61 years	5 (17)

Table 2. Technology believed to play a role in supporting aging in place according to stakeholder groups, categorized in application domains as proposed in the gerontechnology taxonomy.

Application domains	Technologies	O[a]	C[b]	M[c]	T[d]	P[e]
Health and self-esteem						
	Health monitoring	X[f]	X	X	X	X
	Personal alarms	X	X	X	-[g]	X
	Physical activity stimulation	X	-	-	X	X
	Fall detection	-	X	X	-	X
	Medication reminders	-	-	X	X	X
	Wandering detection	-	-	X	X	-
	Online questionnaires	X	-	-	-	X
	Lifestyle monitoring	-	-	-	X	-
Housing and daily living						
	Assistive technology	X	X	X	X	X
	Home automation	X	X	X	X	X
	Household appliances	X	X	X	X	X
	ADL[h]robots	X	-	X	X	X
	Electronic agendas	X	-	-	-	X
	Home adaptations	-	X	-	X	-
	Lift assist devices	-	-	-	X	-
Communication and governance						
	Computers	X	X	X	X	X
	Video telephony	X	X	X	X	X
	Caregiver e-collaboration	X	-	X	X	X
	Electronic health records	X	-	X	-	-
	Social media	-	-	X	-	X
	Telephones	X	-	X	-	-
Work and leisure						
	Television and radio	X	-	X	-	X
	E-readers	X	-	-	X	-
	Games	-	-	-	-	X
Mobility and transport						
	Transportation devices	X	-	X	X	-
	GPS[i]navigation	-	-	-	X	-

[a]O: older adults.

[b]C: care professionals.

[c]M: managers.

[d]T: technologists.

[e]P: policy advisors and policy makers.

[f]X: mentioned by stakeholder group.

[g]-: not mentioned by stakeholder group.

[h]ADL: activities of daily living.

[i]GPS: global positioning system.

Table 3. Stakeholders' perspectives on what constitutes successful implementation of technology for aging in place: major themes, subthemes, and typical quotations.

Major themes	Subthemes	Illustrative quotations	O[a]	C[b]	M[c]	T[d]	P[e]
User-centeredness: Older adults' needs and wishes are given priority during development and deployment of the technology, meaning...							
	...the technology is in accordance with each older adult's specific needs.	"What's needed is a solution for what the individual thinks is a problem, not what we consider a problem" (P #4)	X[f]	X	X	X	X
	...older adults are in control.	"So that it's not the technology that controls my life, but rather it's me controlling the technology" (O #6)	X	X	X	-[g]	-
	...older adults' privacy is treated with respect.	"Seniors shouldn't get the feeling they're being followed or watched" (C #6)	X	X	-	X	-
Acceptance: The technology is accepted by older adults, meaning...							
	...older adults enjoy using the technology.	"A positive experience, causing people to use it again" (M #1)	X	X	X	X	X
	...the technology is used on a regular basis.	"When technology is actually being used" (P #3)	-	X	X	X	X
	...older adults are proud to use the technology (instead of ashamed).	"It shouldn't be stigmatizing" (O #6); "I feel we should aim to create a hype" (M #4)	X	-	X	X	X
Benefits: Use of the technology provides benefits to older adults, meaning...							
	...the technology improves the quality of life of older adults.	"When the client or individual experiences that his or her quality of life remains the same or increases markedly" (M #5)	X	X	X	-	X
	...the technology supports independent living.	"If no one needs to go to a nursing home" (T #2)	-	X	X	X	-
	...the technology provides reassurance.	"Causing people to find an answer to a slowly rising fear of being unstable, frail" (T #5)	X	X	-	X	-
Prerequisites: Favorable prerequisites for ownership and use of technology by older adults exist, meaning...							
	...the technology is easy to use.	"The technology must be extremely user friendly" (M #2)	X	X	X	X	X
	...the technology is affordable.	"Affordability continues to be a problem" (T #6)	X	X	X	X	X
	...the technology is reliable.	"It must work, it must be reliable" (O #3)	X	X	-	X	-
	...technical support is available.	"The supplier or care organization must provide good service" (O #3)	X	X	-	-	X

[a]O: older adults.

[b]C: care professionals.

[c]M: managers.

[d]T: technologists.

[e]P: policy advisors and policy makers.

[f]X: mentioned by stakeholder group.

[g]-: not mentioned by stakeholder group.

Types of Technology That Could Support Aging in Place

Stakeholders had a broad view with regard to technology that could support aging in place, which in their eyes included hardware, software, or combinations of both. In addition, technologies that are not based on ICT were mentioned (eg, consumer appliances and home adaptations). The technologies that were mentioned can be classified in application domains that are part of the gerontechnology taxonomy [27]: health and self-esteem, housing and daily living, mobility and transport, communication and governance, and work and leisure (see Table 2).

In total, 26 different technologies were mentioned by stakeholders across the five domains of the gerontechnology taxonomy. These technologies for the most part fall under the domains of health and self-esteem (8/26, 31%), housing and daily living (7/26, 27%), and communication and governance (6/26, 23%). Out of the 26 technologies mentioned, 5 (19%) fall under the domains of work and leisure or mobility and

transport. Care professionals in total mentioned 9 out of 26 (35%) different types of technology, while the other stakeholder groups each mentioned 17 out of 26 (65%) different types. Out of the 26 technologies, 6 (23%) were mentioned by all stakeholder groups—health monitoring, assistive technology, home automation, household appliances, computers, and video telephony—while 3 (12%) technologies—lifestyle monitoring, lift assist devices, and global positioning system (GPS) navigation—were mentioned by one stakeholder group, the technologists. All other technologies were mentioned by two, three, or four stakeholder groups.

Opinions on What Constitutes a Successful Implementation of Technology

All stakeholder groups considered the implementation of technology for aging in place a success when older adults' needs and wishes are prioritized during development and deployment of technology, the technology is accepted by older adults, the technology provides benefits to older adults, and favorable prerequisites for the use of technology by older adults exist (see Table 3). According to the participants, the aforementioned four major themes—user-centeredness, acceptance, benefits, and prerequisites—are interrelated. All stakeholder groups stressed the importance of taking the perspective of older adults into account, and there was a shared belief that such a user-centered approach would have a positive effect on the acceptance of technology, on the benefits technology can provide, and on the existence of favorable conditions for technology use. Moreover, there was a common belief that technology can only provide benefits to older adults when it is accepted by them, and that acceptance of technology is dependent on certain prerequisites that need to be in place. A typical example of this notion is the following quotation: "Low ease of use leads to nonuse and a lack of added value" (Policy maker/policy advisor #5).

Looking at the first major theme, *user-centeredness*, and its underlying subthemes, all stakeholder groups found it important that technology is in line with the needs of each specific older individual. For example, older adults and policy makers mentioned that technology should not stand in the way of human contact. User-centeredness was also reflected in the fact that stakeholders mentioned that older adults need to be in control over technology instead of the other way around, and that the privacy of older adults needs to be treated with respect. However, policy advisors, care professionals, and older adults also stated that individual differences can make it difficult or expensive for technology to meet older adults' needs in every situation: "It's very hard to achieve this technically...how many diseases are there, and how many different impairments? Think about it" (Older adult #4).

The second major theme, *acceptance*, implicates that older adults enjoy using the technology, and that they use it on a regular basis. It also means that older adults are proud to use technology. The latter point reveals a difference of tone between stakeholder groups. Older adults stressed the importance of not feeling ashamed or stigmatized, while managers, technologists, and policy advisors talked in terms of taking pride in the technology: "It's okay to have it in your home and show it to visitors: 'Look what I have!'...It's not all bad when you grow older, of course you want to show off the nice things that you have" (Technologist #3).

With regard to the third major theme, *benefits*, and its underlying subthemes, stakeholders felt that technology needs to improve older adults' quality of life, support their ability to live independently, and provide reassurance (ie, enhance safety). However, care professionals, managers, and policy advisors stressed that other stakeholder groups are also involved in using technology for aging in place: "People often look at older adults as being the end user. However, informal and professional caregivers are also end users" (Policy advisor/policy maker #2). According to managers, this implies that professional caregivers need to see the benefits of employing technology as well. Older adults felt that technology should provide benefits, but also that technology should not make life too easy: "I think that technology should not make people lazy. For instance, mobility scooters—with all due respect for people who need them—are being used too easily, causing people to walk less" (Older adult #6).

The fourth major theme, *prerequisites*, entails the existence of conditions favorable to technology use and ownership. More specifically, stakeholders mentioned that technology should be easy to use, affordable, and reliable. Additionally, technical support should be available, preferably in person: "I think that there should be a physical location where one can ask something...personal support" (Policy advisor/policy maker #5). Care professionals and technologists, especially, expressed concerns with regard to affordability. Care professionals mentioned that technology in care settings can be expensive, and they worry about who would pay for the technology. Technologists mentioned that they foresee a trend where older adults themselves are the ones who pay for technology. In this scenario, technologists see older adults' willingness to pay for technology as critical, and they feel that the technology they wish to sell needs to be more affordable than competing alternatives. In contrast, older adults only fleetingly mentioned the fact that technology needs to be affordable. As for managers, they looked at affordability from a cost-benefit perspective: "When the financial benefits exceed the investments" (Manager #1).

What is Needed to Successfully Implement Technology for Aging in Place

Looking at their own roles, stakeholders mentioned several things that they need or can contribute to enable successful implementations of technology for aging in place. These can be organized into four major themes and eight underlying subthemes (see Table 4).

Table 4. Stakeholders' views on what is needed to successfully implement technology for aging in place: major themes and subthemes.

Major themes	Subthemes	O[a]	C[b]	M[c]	T[d]	P[e]
Take the leap						
	Change in attitude(s)	X[f]	X	X	X	-[g]
	Change in policies	-	X	X	X	X
	Collaborate with other organizations	-	-	X	-	X
Bridge the gap						
	Match technology with individuals	-	X	X	-	X
	Stimulate interdisciplinary education	-	-	-	-	X
Facilitate technology for the masses						
	Work on products and research that support large-scale rollouts	-	-	X	X	-
	Train target groups on how to use technology	X	X	-	-	-
Take time to reflect	Evaluate use and outcomes	-	X	-	-	X

[a]O: older adults.

[b]C: care professionals.

[c]M: managers.

[d]T: technologists.

[e]P: policy advisors and policy makers.

[f]X: mentioned by stakeholder group.

[g]-: not mentioned by stakeholder group.

The first theme, *take the leap*, is concerned with what is needed in terms of commitment by stakeholders. Most stakeholder groups emphasized that a change in attitude is needed on their part to achieve successful implementations. For example, older adults mentioned that they could be more assertive. By this, it was meant that older adults can improve in "Saying what you think, desire, and feel" (Older adult #5), and that older adults are prepared to ask for help. Older adults stated that this is particularly important when talking to technologists. Additionally, older adults mentioned that they sometimes need to be stimulated to use technology, or as one older adult phrased it, "Pushed gently" (Older adult #6). Reflecting on their own role, care professionals stated that they need to adjust, and accept that things are changing: "From a caring perspective, I want to help people in person...however, some things are no longer feasible. I feel that a new mindset is needed" (Care professional #7) and "It's the client who has technology in his home, and we need to become accustomed to it" (Care professional #4). Managers felt that they need to promote the use of technology more. They mentioned that they could initiate pilot projects, which are seen as a way to have care professionals gain experience in using technology. Technologists mentioned that technology companies need to be prepared to take financial risks. More specifically, companies need to have the confidence to produce and roll out technologies on a large scale. For this, a long-term strategy and perseverance are required: "There can be up to 20 years between designing the thing, and starting to make a profit. We have to get used to that; that's the long-term vision we have to have" (Technologist #3).

Additionally, most stakeholder groups proposed that policies need to be changed. Care professionals ask that the organizations that they work for formulate a privacy policy for situations in which technology is employed. Managers stated that they would like more flexibility with regard to the relevant laws and regulations. They also mentioned that they need to incorporate technology in their organizational strategy: "It all starts at the top: what are the priorities for the organization in the years to come? When technology isn't in there..." (Manager #5). Reflecting on their own roles, policy advisors and policy makers mentioned that a large proportion of technology for older adults is being subsidized, and that the use of these technologies is frequently not sustainable: "When the funding stops...the technology is no longer used" (Policy advisor/policy maker #2). They argue that they need to find ways to counter this unwanted effect of current policies. Some technologists noted that subsidizing technology may obscure the actual needs of potential clients: "When people receive something for free, I can't make out whether they actually want it" (Technologist #1).

Furthermore, the need for more organizational collaboration was mentioned by managers and policy advisors. Managers within home care or social work organizations felt a need to collaborate with others outside of their own organization in order to enable successful implementations of technology for aging in place: "I can't do it alone. I need the municipality, and collaboration with the housing association and welfare organizations. You have to combine forces" (Manager #4). In this respect, insurance companies, patients' associations, and informal caregivers were also mentioned. Policy advisors and policy makers emphasized the importance of international and interdisciplinary collaboration.

The second theme, *bridge the gap*, entails the work that is needed to connect available technological solutions to the needs

of each specific older adult. Care professionals, managers, and policy makers stated that help is needed to be able to match technology with individuals. Care professionals mentioned that they would benefit from a "decision tool." Such a tool should allow care professionals to find and select the appropriate technology or combination of technologies for each specific client. Ideally, the technologies and aids that are deployed should also be registered in electronic health records. The managers in the study, who worked for different organizations than the care professionals, also mentioned that they would like to provide the care professionals with such a decision tool. Moreover, managers stated they would like to work together with a person (ie, consultant) who knows which technologies are on the market, and who can match these with the problems older adults face while trying to maintain their independence. Policy makers and policy advisors felt that interdisciplinary education is required to achieve this: “Because you need to know what an individual needs, you have to understand that person, and subsequently you have to know how to arrange technologies, services, and care” (Policy advisor/policy maker #3).

With regard to the third theme, *facilitate technology for the masses*, managers and technologists discussed the need to engage in large-scale rollouts of technology. Managers stated that there is a demand for technological solutions that can benefit a large proportion of older adults. In their eyes, large-scale rollouts can increase the willingness of commercial companies to invest, which is seen as a requirement for making technology for aging in place affordable. In their perception, more research is needed to provide scientific evidence that technology for aging in place is effective, and this is also expected to increase support by the government. To be able to conduct large-scale rollouts, technologists mentioned that companies need to do more research in order to gain a more profound understanding of what drives or impedes technology use by older adults.

Additionally, comments were made with regard to empowering target groups to be able to take advantage of technology. Older adults stated that they need to attend courses to learn how to use technology when they are still healthy enough to attend them. Care professionals also mentioned that they need training to be able to work with the technology. In their eyes, this applies to inexperienced as well as experienced care professionals: “You have to let yourself get educated, particularly those of us who have been working for a long time” (Care professional #2).

The last theme, *take time to reflect*, entails the evaluation of use and outcomes. Care professionals mentioned that they see it as their responsibility to regularly evaluate whether the use of technology is appropriate and not too excessive: “You shouldn’t use technology for everything” (Care professional #5). Additionally, policy makers stated that they feel a need to measure whether the use of technology is successful in terms of the desired outcomes. They see it as their role to promote evidence-based solutions.

Discussion

Principal Findings

This study aimed to understand the positions of stakeholders who are involved in the implementation of technology for aging in place—older adults, care professionals, managers of care organizations, technologists, and policy makers. It was found that stakeholders considered a multitude of technologies to be relevant for enabling independent living. However, it is important to note that only a small number of technologies were mentioned by all stakeholder groups. Furthermore, care professionals mentioned considerably fewer different types of technology than other stakeholder groups, which is in line with previous research [8]. Additionally, studies have shown that older adults may not be aware of technologies that could be of benefit to them [28,29]. Therefore, when planning and initiating projects concerned with technological solutions for aging in place, it is advisable to take into account that stakeholders may have a limited understanding of the scope of available technologies, and that stakeholders may differ in their awareness of available technologies. Moreover, technologies that are not ICT based (eg, household appliances and home adaptations) are also relevant in the context of aging in place according to stakeholders. In this sense, their concept of technology is less exclusive than the commonly used definitions of ambient assisted living technology [30], smart home technology [6], and eHealth [31].

With regard to the aims of stakeholders, all stakeholder groups felt that the implementation of technology for aging in place can be considered a success when (1) older adults’ needs and wishes are prioritized during development and deployment of the technology, (2) the technology is accepted by older adults, (3) the technology provides benefits to older adults, and (4) favorable prerequisites for the use of technology by older adults exist. As such, all stakeholder groups were profoundly concerned with the position of older adults when it comes to implementing technologies for aging in place. This study aligns closely with work reported by Greenhalgh et al [32], in which the authors sought to define quality in the design, implementation, and use of telehealth and telecare solutions for older adults with assisted living needs. In that study, which involved older adults, technology suppliers, and service providers, it was concluded that every stakeholder needs to comprehend the (changing) needs and capabilities of older adults, as well as their social context [32]. Such an approach, centered around the older individual, also aligns with the trend toward patient empowerment and patient engagement [33-36]; technology may be used to empower seniors, but this requires their engagement during design and implementation.

While the stakeholders in this study generally appeared to have identical aims with regard to technology for aging in place, it is important to note that underlying differences existed between stakeholders. For example, all stakeholder groups agreed that technology should provide certain benefits to older adults, but older adults were the only group that stressed that technology should not provide too many benefits, since this could make people dependent on technology, which is in line with previous

research [12,37,38]. Another example of the variance of opinion is affordability: stakeholders agreed that this is important, but they did not seem to be on the same page with regard to who should pay for the technology. Participants in this study were not involved in a joint effort to implement technology at the time that data for this study was gathered. Once stakeholders are further into the process of implementing technology together, the aforementioned differences in the interpretation of key aims, such as benefits and affordability, could lead to cases of *stakeholder dissonance*, which threatens a project's viability if left undetected and unresolved [39].

Each stakeholder group mentioned specific steps that need to be taken to achieve successful implementations. Collectively, stakeholders felt they need to take the leap (ie, change attitudes, change policies, and collaborate with other organizations), bridge the gap (ie, match technology with individuals and stimulate interdisciplinary education), facilitate technology for the masses (ie, work on products and research that supports large-scale rollouts and train target groups on how to use technology), and take time to reflect (ie, evaluate use and outcomes). Some of the aforementioned steps or recommended actions have also been reported by similar stakeholder groups in other studies; for example, the need to focus on changing the attitudes of care receivers and care givers [40,41], the need to match technology with individual clients [28,40,42], and the need for training stakeholder groups [8,42,43]. Additionally, studies have pointed to recommended actions that were not mentioned by participants in this study. These include the need to consider how the introduction of technology affects the existing workflow in home care organizations [40-42] and the fact that care professionals require support while using technology [8,44,45].

The recommended actions brought forward by stakeholders in this study imply that structural changes need to be made on political/ strategic, organizational/ contractual, managerial/ scientific, and operational levels [46]. Such changes will not be easy to implement because of their fundamental character, and because they require changes in how different stakeholder groups operate and interface with one another [11,22,32]. Additionally, recent evaluations of the Delivering Assisted Living Lifestyles at Scale (dallas) program in England [11] and Scotland [22] indicate that while involving end users in the design of technologies could promote adoption, it is also very difficult to simultaneously codesign and deliver technologies at a large scale. The reason for this is that codesign is time and resource consuming [11,22]. This was also demonstrated by Linskell and Bouamrane [47], who described two possible routes for the delivery of technology that could support aging in place: a short and direct delivery route, which is prone to misinterpretation of user needs, and a longer codesign route which incorporates task analysis and more extensive specification of product requirements. Therefore, when it comes to matching technology with individuals, the challenge seems to lie in being able to determine when a short and direct delivery route is acceptable and when a longer codesign route is warranted.

The results of this study can be viewed in light of the Normalization Process Theory (NPT), as described by May and Finch [48-50]. NPT addresses "the factors needed for successful implementation and integration of interventions into routine work" [49] and consists of four main components: coherence (ie, meaning and sense making by stakeholders); cognitive participation (ie, commitment and engagement by stakeholders); collective action (ie, the work stakeholders do to make the intervention function in practice); and reflexive monitoring (ie, formal and informal appraisal of the benefits and costs of the intervention) [50]. Our findings seem to indicate that NPT can potentially provide a useful framework for studying implementations in the context of aging in place. First, the themes that emerged in this study with regard to what is needed to successfully implement technology for aging in place bear resemblance to NPT's concepts of cognitive participation, collective action, and reflexive monitoring. For example, the *take the leap* theme, which includes a change in attitudes, a change in policies, and collaboration with other organizations, resembles NPT's cognitive participation component; the *bridge the gap* and *facilitate technology for the masses* themes are in line with NPT's component of collective action. Second, NPT's first component, coherence, includes a "shared understanding of the aims, objectives, and expected benefits" [51], and this study shows that focus group sessions can be employed to start to develop this type of shared understanding. However, it was not our goal to verify or test NPT in this study. Future studies are necessary to explore the value of NPT in the context of aging in place, particularly in situations where available technological solutions need to be matched to the specific needs of each client. Furthermore, focus group sessions in this study were mono-disciplinary and led to findings that pointed to several differences among stakeholder groups, indicating that it would be beneficial to follow up on these mono-disciplinary sessions by conducting heterogeneous sessions to further develop coherence.

Limitations

Our study is limited by the fact that it may not have included all the relevant stakeholders. For example, research shows that family members and informal caregivers can play an important role in the (effective) use of technology by community-dwelling older adults [38,52]. Additionally, the grouping of stakeholders in this study is an oversimplification, as each stakeholder group can be broken down into more specific subgroups. Furthermore, process evaluations covering a longer period of time are needed to determine how dynamics between stakeholders influence the effective provisioning of personalized and appropriate technology that can help older adults to age in place. Lastly, it cannot be ruled out that our study was susceptible to selection bias since all participants were part of a project that aimed to improve the deployment of technology for aging in place by conducting research in the homes of older adults.

Conclusions

In conclusion, this study adds to the limited body of work concerned with successfully implementing technology that aims to support aging in place. Stakeholders in this study largely agree on the direction in which they should be heading, yet they have different perspectives with regard to the technologies that can be employed and the work that needs to be done to

implement these. Central to a successful implementation seems to be the tailoring of technology or technologies to the specific needs of each community-dwelling individual, and the work that is needed by stakeholders to support this type of service delivery on a large scale. Our findings indicate a tension between aiming to personalize technology implementations and aiming to deploy technology en masse. It is clear that, in order to successfully implement technology for aging in place, stakeholders need to engage in an ongoing mutual commitment focused on the goal of empowering older adults through the use of technology.

Acknowledgments

The authors would like to thank the participants of the focus groups. Additionally, Marianne Nieboer, Babs Bergs, Rienk Overdiep, Sil Aarts, Joost van Hoof, Claire van der Voort, Stephan Rooijers, Caroline Heesakkers, Jac Thielemans, Erik van Santvoort, and Anna van der Straaten are thanked for their help in moderating and organizing the sessions. The authors would also like to thank Kimberley Abeln, Tamara de Bruin, Femke Dilven, Ger Gijsbers, and Suzanne van Gijzel for transcribing the interviews and for their help during data analysis.

This work was supported by the Regional Attention and Action for Knowledge Circulation (RAAK) scheme (PRO-3-37), which is managed by the Foundation Innovation Alliance (Stichting Innovatie Alliantie [SIA]) with funding from the Dutch Ministry of Education, Culture and Science (OCW). SIA-RAAK had no role in the study design; the collection, analysis, and interpretation of data; the writing of the report; nor the decision to submit the paper for publication.

Authors' Contributions

All authors have made a substantial, direct, intellectual contribution to this study. STMP was responsible for the study concept and design, data analysis, and drafting of the manuscript. EJMW was responsible for the study design and critical revision of the manuscript for important intellectual content. KGL and HJMV were responsible for the critical revision of the manuscript for important intellectual content. All authors approved the final version.

Conflicts of Interest

None declared.

References

1. World Population Ageing 2013. New York, NY: United Nations, Department of Economic and Social Affairs, Population Division; 2013. URL: http://www.un.org/en/development/desa/population/publications/pdf/ageing/WorldPopulationAgeingReport2013.pdf [accessed 2016-04-14] [WebCite Cache ID 6gleQDoHp]
2. Vasunilashorn S, Steinman BA, Liebig PS, Pynoos J. Aging in place: Evolution of a research topic whose time has come. J Aging Res 2012;2012:120952 [FREE Full text] [doi: 10.1155/2012/120952] [Medline: 22175020]
3. Peek STM, Wouters EJM, van Hoof J, Luijkx KG, Boeije HR, Vrijhoef HJM. Factors influencing acceptance of technology for aging in place: A systematic review. Int J Med Inform 2014 Apr;83(4):235-248 [FREE Full text] [doi: 10.1016/j.ijmedinf.2014.01.004] [Medline: 24529817]
4. Michel JP, Franco A. Geriatricians and technology. J Am Med Dir Assoc 2014 Dec;15(12):860-862. [doi: 10.1016/j.jamda.2014.09.016] [Medline: 25454316]
5. Wilson C, Hargreaves T, Hauxwell-Baldwin R. Smart homes and their users: A systematic analysis and key challenges. Pers Ubiquitous Comput 2015 Feb;19(2):463-476. [doi: 10.1007/s00779-014-0813-0]
6. Balta-Ozkan N, Davidson R, Bicket M, Whitmarsh L. Social barriers to the adoption of smart homes. Energy Policy 2013;63:363-374.
7. Doyle J, Bailey C, Ni Scanaill C, van den Berg F. Lessons learned in deploying independent living technologies to older adults' homes. Univers Access Inf Soc 2014 Jun;13(2):191-204. [doi: 10.1007/s10209-013-0308-1]
8. Nieboer ME, van Hoof J, van Hout AM, Aarts S, Wouters EJM. Professional values, technology and future health care: The view of health care professionals in The Netherlands. Technol Soc 2014 Nov;39:10-17. [doi: 10.1016/j.techsoc.2014.05.003]
9. Ehrenhard M, Kijl B, Nieuwenhuis L. Market adoption barriers of multi-stakeholder technology: Smart homes for the aging population. Technol Forecast Soc Change 2014 Nov;89:306-315. [doi: 10.1016/j.techfore.2014.08.002]
10. Sponselee AAG. Acceptance and effectiveness of smart home solutions. In: van Hoof J, Demiris G, Wouters EJM, editors. Handbook of Smart Homes, Health Care and Well-Being. Cham, Switzerland: Springer International Publishing; Nov 4, 2014:1-12.
11. Devlin AM, McGee-Lennon M, O'Donnell CA, Bouamrane MM, Agbakoba R, O'Connor S, “dallas” evaluation team. Delivering digital health and well-being at scale: Lessons learned during the implementation of the dallas program in the United Kingdom. J Am Med Inform Assoc 2016 Jan;23(1):48-59 [FREE Full text] [doi: 10.1093/jamia/ocv097] [Medline: 26254480]

12. Peek STM, Aarts S, Wouters EJM. Can smart home technology deliver on the promise of independent living? A critical reflection based on the perspectives of older adults. In: van Hoof J, Demiris G, Wouters EJM, editors. Handbook of Smart Homes, Health Care and Well-Being. Cham, Switzerland: Springer International Publishing; Mar 6, 2015:1-10.
13. Reeder B, Meyer E, Lazar A, Chaudhuri S, Thompson HJ, Demiris G. Framing the evidence for health smart homes and home-based consumer health technologies as a public health intervention for independent aging: A systematic review. Int J Med Inform 2013 Jul;82(7):565-579. [doi: 10.1016/j.ijmedinf.2013.03.007] [Medline: 23639263]
14. Graybill EM, McMeekin P, Wildman J. Can aging in place be cost effective? A systematic review. PLoS One 2014;9(7):e102705 [FREE Full text] [doi: 10.1371/journal.pone.0102705] [Medline: 25058505]
15. Boonstra A, van Offenbeek M. Towards consistent modes of e-health implementation: Structurational analysis of a telecare programme's limited success. Inf Syst J 2010 Aug 27;20(6):537-561. [doi: 10.1111/j.1365-2575.2010.00358.x]
16. Freeman RE. Strategic Management: A Stakeholder Approach. Boston, MA: Pitman Publishing Inc; Mar 2010:1-292.
17. Frooman J. Stakeholder influence strategies. Acad Manage Rev 1999 Apr 01;24(2):191-205. [doi: 10.5465/AMR.1999.1893928]
18. Clark J, McGee Lennon M. A stakeholder-centred exploration of the current barriers to the uptake of home care technology in the UK. J Assist Technol 2011 Mar 18;5(1):12-25. [doi: 10.5042/jat.2011.0097]
19. Reginatto BMB. Understanding barriers to wider telehealth adoption in the home environment of older people: An exploratory study in the Irish context. J Adv Life Sci 2012;4(3&4):63-76.
20. Leonard KJ. Critical success factors relating to healthcare's adoption of new technology: A guide to increasing the likelihood of successful implementation. Electron Healthc 2004 Mar;2(4):72-81.
21. Sponselee A, Schouten B, Bouwhuis D, Willems C. Smart home technology for the elderly: Perceptions of multidisciplinary stakeholders. In: Mühlhäuser M, Ferscha A, Aitenbichler E, editors. Constructing Ambient Intelligence. Berlin, Germany: Springer; 2008:314-326.
22. Agbakoba R, McGee-Lennon M, Bouamrane M, Watson N, Mair FS. Implementation factors affecting the large-scale deployment of digital health and well-being technologies: A qualitative study of the initial phases of the 'Living-It-Up' programme. Health Informatics J 2015 Aug 14:1-11. [doi: 10.1177/1460458215594651] [Medline: 26276795]
23. Sugarhood P, Wherton J, Procter R, Hinder S, Greenhalgh T. Technology as system innovation: A key informant interview study of the application of the diffusion of innovation model to telecare. Disabil Rehabil Assist Technol 2014 Jan;9(1):79-87. [doi: 10.3109/17483107.2013.823573] [Medline: 23919407]
24. Murray E, Burns J, May C, Finch T, O'Donnell C, Wallace P, et al. Why is it difficult to implement e-health initiatives? A qualitative study. Implement Sci 2011;6:6 [FREE Full text] [doi: 10.1186/1748-5908-6-6] [Medline: 21244714]
25. Sim J. Collecting and analysing qualitative data: Issues raised by the focus group. J Adv Nurs 1998 Aug;28(2):345-352. [Medline: 9725732]
26. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
27. van Bronswijk JEMH, Bouma H, Fozard JL. Technology for quality of life: An enriched taxonomy. Gerontechnology 2002;2(2):169-172. [doi: 10.4017/gt.2002.02.02.001.00]
28. Jarrold D, Bouamrane MM. Engaging with end-users of independent living technology in Scotland... The experience of the Blackwood Foundation. In: Proceedings of the 5th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth) and Workshops. New York, NY: IEEE; 2011 Presented at: 5th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth) and Workshops; May 23-26, 2011; Dublin, Ireland p. 278-282. [doi: 10.4108/icst.pervasivehealth.2011.246045]
29. Heinz M, Martin P, Margrett JA, Yearns M, Franke W, Yang H, et al. Perceptions of technology among older adults. J Gerontol Nurs 2013 Jan;39(1):42-51. [doi: 10.3928/00989134-20121204-04] [Medline: 23244061]
30. Pecchia L, Chen LL, Nugent C, Bravo J, editors. Preface. In: Ambient Assisted Living and Daily Activities. Cham, Switzerland: Springer International Publishing; Dec 2014.
31. Pagliari C, Sloan D, Gregor P, Sullivan F, Detmer D, Kahan JP, et al. What is eHealth (4): A scoping exercise to map the field. J Med Internet Res 2005;7(1):e9 [FREE Full text] [doi: 10.2196/jmir.7.1.e9] [Medline: 15829481]
32. Greenhalgh T, Procter R, Wherton J, Sugarhood P, Hinder S, Rouncefield M. What is quality in assisted living technology? The ARCHIE framework for effective telehealth and telecare services. BMC Med 2015;13(91):1-15 [FREE Full text] [doi: 10.1186/s12916-015-0279-6] [Medline: 25902803]
33. Calvillo J, Román I, Roa LM. How technology is empowering patients? A literature review. Health Expect 2015 Oct;18(5):643-652. [doi: 10.1111/hex.12089] [Medline: 23711169]
34. Vrijhoef HJM. Patient engagement is the new hard currency in health care. Int J Care Coord 2015 Mar;18(1):3-4. [doi: 10.1177/2053434515577291]
35. Richards T, Montori VM, Godlee F, Lapsley P, Paul D. Let the patient revolution begin. BMJ 2013;346:f2614. [Medline: 23674136]
36. Engelen L. BMJ Blogs. 2013 Aug 16. Patients not included URL: http://blogs.bmj.com/bmj/2013/08/16/lucien-engelen-patients-not-included/ [accessed 2015-10-16] [WebCite Cache ID 6cK47IXlD]

37. Bowman S, Hooker K, Steggell CD, Brandt J. Perceptions of communication and monitoring technologies among older rural women: Problem or panacea? J Hous Elderly 2013 Jan;27(1-2):48-60. [doi: 10.1080/02763893.2012.754814]
38. Peek STM, Luijkx KG, Rijnaard MD, Nieboer ME, van der Voort CS, Aarts S, et al. Older adults' reasons for using technology while aging in place. Gerontology 2016 Feb;62:226-237 [FREE Full text] [doi: 10.1159/000430949] [Medline: 26044243]
39. Samaras E, Samaras G. Stakeholder dissonance as a critical determinant of an e-health initiative: A case study. Online J Nurs Inform 2012;16(1):36-38.
40. Koopman RJ, Wakefield BJ, Johanning JL, Keplinger LE, Kruse RL, Bomar M, et al. Implementing home blood glucose and blood pressure telemonitoring in primary care practices for patients with diabetes: Lessons learned. Telemed J E Health 2014 Mar;20(3):253-260 [FREE Full text] [doi: 10.1089/tmj.2013.0188] [Medline: 24350806]
41. Essén A, Conrick M. New e-service development in the homecare sector: Beyond implementing a radical technology. Int J Med Inform 2008 Oct;77(10):679-688. [doi: 10.1016/j.ijmedinf.2008.02.001] [Medline: 18514021]
42. Hebert MA, Korabek B. Stakeholder readiness for telehomecare: Implications for implementation. Telemed J E Health 2004;10(1):85-92. [doi: 10.1089/153056204773644625] [Medline: 15104920]
43. Agbakoba R, McGee-Lennon M, Bouamrane MM, Watson N, Mair F. Implementing a national Scottish digital health & wellbeing service at scale: A qualitative study of stakeholders' views. Stud Health Technol Inform 2015;216:487-491. [Medline: 26262098]
44. de Veer AJE, Fleuren MAH, Bekkema N, Francke AL. Successful implementation of new technologies in nursing care: A questionnaire survey of nurse-users. BMC Med Inform Decis Mak 2011 Oct 27;11(67):1-12 [FREE Full text] [doi: 10.1186/1472-6947-11-67] [Medline: 22032728]
45. Robben SH, Perry M, Huisjes M, van Nieuwenhuijzen L, Schers HJ, van Weel C, et al. Implementation of an innovative web-based conference table for community-dwelling frail older people, their informal caregivers and professionals: A process evaluation. BMC Health Serv Res 2012 Aug 15;12(251):1-12 [FREE Full text] [doi: 10.1186/1472-6963-12-251] [Medline: 22894654]
46. Vitacca M, Mazzù M, Scalvini S. Socio-technical and organizational challenges to wider e-Health implementation. Chron Respir Dis 2009;6(2):91-97. [doi: 10.1177/1479972309102805] [Medline: 19411570]
47. Linskell J, Bouamrane MM. Assisted-living spaces for end-users with complex needs: A proposed implementation and delivery model. Health Informatics J 2012 Sep;18(3):159-170. [doi: 10.1177/1460458212441474] [Medline: 23011811]
48. May C, Finch T. Implementing, embedding, and integrating practices: An outline of normalization process theory. Sociology 2009 Jun 15;43(3):535-554. [doi: 10.1177/0038038509103208]
49. 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):1-9 [FREE Full text] [doi: 10.1186/1748-5908-4-29] [Medline: 19460163]
50. Murray E, Treweek S, Pope C, MacFarlane A, Ballini L, Dowrick C, et al. Normalisation process theory: A framework for developing, evaluating and implementing complex interventions. BMC Med 2010;8(63):1-11 [FREE Full text] [doi: 10.1186/1741-7015-8-63] [Medline: 20961442]
51. Mair FS, May C, O'Donnell C, Finch T, Sullivan F, Murray E. Factors that promote or inhibit the implementation of e-health systems: An explanatory systematic review. Bull World Health Organ 2012 May 1;90(5):357-364 [FREE Full text] [doi: 10.2471/BLT.11.099424] [Medline: 22589569]
52. Luijkx K, Peek S, Wouters E. "Grandma, you should do it--It's cool": Older adults and the role of family members in their acceptance of technology. Int J Environ Res Public Health 2015;12(12):15470-15485 [FREE Full text] [doi: 10.3390/ijerph121214999] [Medline: 26690188]

Abbreviations

ADL: activities of daily living
C: care professionals
dallas: Delivering Assisted Living Lifestyles at Scale
GPS: global positioning system
ICT: information and communications technology
M: managers
NPT: Normalization Process Theory
O: older adults
OCW: Dutch Ministry of Education, Culture and Science
P: policy advisors and policy makers
RAAK: Regional Attention and Action for Knowledge Circulation
SIA: Stichting Innovatie Alliantie
T: technologists

Edited by G Eysenbach; submitted 16.10.15; peer-reviewed by MM Bouamrane, J Ross, S Hinder; comments to author 12.11.15; revised version received 17.01.16; accepted 04.02.16; published 03.05.16.

Please cite as:
Peek STM, Wouters EJM, Luijkx KG, Vrijhoef HJM
What it Takes to Successfully Implement Technology for Aging in Place: Focus Groups With Stakeholders
J Med Internet Res 2016;18(5):e98
URL: http://www.jmir.org/2016/5/e98/
doi:10.2196/jmir.5253
PMID:27143097

Original Paper

Preexposure Prophylaxis Modality Preferences Among Men Who Have Sex With Men and Use Social Media in the United States

Eric William Hall[1], MPH; Walid Heneine[2], PhD; Travis Sanchez[1], MPH, DVM; Robert Craig Sineath[1], MPH; Patrick Sullivan[1], DVM, PhD

[1]Department of Epidemiology, Rollins School of Public Health, Emory University, Atlanta, GA, United States

[2]National Center for HIV/AIDS, Viral Hepatitis, STD, and TB Prevention, Centers for Disease Control and Prevention, Atlanta, GA, United States

Corresponding Author:
Eric William Hall, MPH
Department of Epidemiology
Rollins School of Public Health
Emory Univeresity
GCR 432
1518 Clifton Road
Atlanta, GA, 30322
United States
Phone: 1 4047278965
Fax: 1 4047128392
Email: eric.w.hall@emory.edu

Abstract

Background: Preexposure prophylaxis (PrEP) is available as a daily pill for preventing infection with the human immunodeficiency virus (HIV). Innovative methods of administering PrEP systemically or topically are being discussed and developed.

Objective: The objective of our study was to assess attitudes toward different experimental modalities of PrEP administration.

Methods: From April to July 2015, we recruited 1106 HIV-negative men who have sex with men through online social media advertisements and surveyed them about their likelihood of using different PrEP modalities. Participants responded to 5-point Likert-scale items indicating how likely they were to use each of the following PrEP modalities: a daily oral pill, on-demand pills, periodic injection, penile gel (either before or after intercourse), rectal gel (before/after), and rectal suppository (before/after). We used Wilcoxon signed rank tests to determine whether the stated likelihood of using any modality differed from daily oral PrEP. Related items were combined to assess differences in likelihood of use based on tissue or time of administration. Participants also ranked their interest in using each modality, and we used the modified Borda count method to determine consensual rankings.

Results: Most participants indicated they would be somewhat likely or very likely to use PrEP as an on-demand pill (685/1105, 61.99%), daily oral pill (528/1036, 50.97%), injection (575/1091, 52.70%), or penile gel (438/755, 58.01% before intercourse; 408/751, 54.33% after). The stated likelihoods of using on-demand pills (median score 4) and of using a penile gel before intercourse (median 4) were both higher than that of using a daily oral pill (median 4, $P<.001$ and $P=.001$, respectively). Compared with a daily oral pill, participants reported a significantly lower likelihood of using any of the 4 rectal modalities (Wilcoxon signed rank test, all $P<.001$). On 10-point Likert scales created by combining application methods, the reported likelihood of using a penile gel (median 7) was higher than that of using a rectal gel (median 6, $P<.001$), which was higher than the likelihood of using a rectal suppository (median 6, $P<.001$). The modified Borda count ranked on-demand pills as the most preferred modality. There was no difference in likelihood of use of PrEP (gel or suppository) before or after intercourse.

Conclusions: Participants typically prefer systemic PrEP and are less likely to use a modality that is administered rectally. Although most of these modalities are seen as favorable or neutral, attitudes may change as information about efficacy and application becomes available. Further data on modality preference across risk groups will better inform PrEP development.

(J Med Internet Res 2016;18(5):e111) doi:10.2196/jmir.5713

KEYWORDS

preexposure prophylaxis; PrEP; men who have sex with men; MSM; HIV prevention; Facebook; survey

Introduction

Although improvements in treatment have extended the life expectancy of people infected with the human immunodeficiency virus (HIV), there are still a troubling number of new HIV infections each year. In particular, HIV incidence rates among men who have sex with men (MSM) are increasing in North America and several other regions of the world [1,2]. In 2014, 70% of all new infections in the United States occurred among MSM [3]. To reduce the number of new infections, prevention strategies targeted toward specific risk groups are needed.

A recent and exciting strategy, HIV preexposure prophylaxis (PrEP), involves using antiretroviral medication to reduce the risk of HIV infection among HIV-negative individuals. In 2010, this concept was first demonstrated in humans with publication of results from the iPrEx study, a randomized controlled trial that tested the efficacy of tenofovir disoproxil fumarate (TDF) in combination with emtricitabine (FTC) among MSM [4]. The iPrEx trial demonstrated a 44% (95% CI 15%-63%) reduction in HIV incidence among men who were taking TDF/FTC as a daily oral pill compared with a placebo group. Since then, other randomized controlled trials have indicated that a daily PrEP pill can reduce risk of HIV transmission in HIV-discordant couples [5], sexually active heterosexual men and women [6], and intravenous drug users [7]. However, 2 other trials showed no protective effect in heterosexual women in high-risk areas of Africa, but both of these studies had problems with adherence [8,9].

Results from these PrEP efficacy studies show varying degrees of HIV risk reduction that ranges from 0% to 75%. However, this wide range is most commonly attributed to varying levels of adherence, because PrEP efficacy is much higher among participants who demonstrated consistent use [10-14]. For example, when blood samples were analyzed in the iPrEx study, HIV incidence reduction was 92% among participants who had the drug detectable in their blood and 99% among participants who had drug levels corresponding to daily use (both compared with the placebo group) [15,16].

In 2014, the US Centers for Disease Control and Prevention and the World Health Organization released guidelines that recommend the use of daily oral PrEP in populations with an elevated risk of HIV infection [17,18]. However, barriers such as cost, the burden of taking a daily pill, and concerns about potential health effects (both long-term and short-term effects) have led to PrEP being underused by eligible people [13,19-21].

Although a daily oral TDF/FTC pill is the only approved and recommended form of PrEP, there is growing interest in developing new methods for administering antiretroviral drugs as prevention. Topical applications of PrEP have been studied in 2 clinical trials of high-risk women who used a TDF-based vaginal gel either before and after sex [22] or on a daily basis [8]. Although the results from these trials present conflicting conclusions, further analysis indicates that the efficacy of vaginal gel depends on the concentration of tenofovir in the cervicovaginal fluid, which is also likely an indication of adherence [23]. A recent study among MSM (iPERGAY) investigated the efficacy of intermittent TDF/FTC pills taken just before and after sexual encounters, but the study was stopped early because initial analysis found comparable protection against HIV infection between these on-demand PrEP regimens and daily PrEP [24].

Researchers are investigating the delivery of PrEP as oral pills that are used intermittently (ie, less than a daily basis), topical gels, vaginal rings, and long-lasting injections [21,25,26]. Because high levels of PrEP efficacy are dependent on adherence, there is obvious interest in developing administration modalities that target groups are willing to use. To help guide these research efforts, we sought to assess attitudes among MSM toward a variety of potential modalities of PrEP administration.

Methods

Recruitment

We collected data through a study that primarily explored alternative methods for delivering consent information and maximizing retention in online surveys [27] (funded by NICHD Research Grant 1R21-HD074502-01A1). Participants were recruited through targeted advertisements on a social media website (Facebook) from April 2015 to July 2015. Recruitment advertisements appeared to users in the United States who indicated on their Facebook profile that they are male, over 18 years of age, and interested in men. People who clicked on the advertisements were directed to an online consent module and a short screener to determine eligibility. To be eligible for the survey, users had to be male, between 18 and 34 years old, and not report having sex only with women in the past year. Men who reported never having oral or anal sex with a man were removed from the analysis dataset.

Eligible men were given an online survey that collected demographic information such as age, education, race or ethnicity, zip code, and self-identified sexual orientation. The survey also collected information about sexual history and current sexual practices, history of HIV testing, and relationship status. All study materials and procedures were approved by the Emory University Institutional Review Board.

Participants who reported a negative or unknown HIV status were asked about their knowledge of PrEP and history of use. Participants who had not previously used PrEP answered 5-point Likert-scale items that asked how likely they were to use different PrEP modalities to reduce the risk of getting HIV. They were asked about 9 Likert-scale items, 1 for each of the modalities listed in Table 1. Participants were only asked about modalities that involved penile application if they reported having insertive anal sex in the past year. Likewise, participants were only asked about the rectal modalities if they reported having receptive anal sex in the past year. We collected responses to each Likert-scale item in the following format: 1=very unlikely, 2=somewhat unlikely, 3=neither likely or unlikely, 4=somewhat likely, 5=very likely.

Table 1. Preexposure prophylaxis modalities presented[a] to online survey respondents aged 18–34 years, by type of anal sex with male partner(s) in the past year, United States, April–July 2015.

Modalities	Insertive only	Receptive only	No anal sex
Daily oral pill	X[b]	X	X
On-demand pills[c]	X	X	X
Injection every 1–3 months	X	X	X
Penis gel before insertive intercourse	X		
Penis gel after insertive intercourse	X		
Rectal gel before receptive intercourse		X	
Rectal gel after receptive intercourse		X	
Rectal suppository 30 minutes before receptive intercourse		X	
Rectal suppository 3 hours after receptive intercourse		X	

[a]The survey included individual Likert-scale items asking the likelihood of using each modality.

[b]Participants who indicated they had both insertive and receptive anal sex in the past year were presented all modalities. Modalities presented to participants who said they only had insertive anal sex, receptive anal sex, or no anal sex in the past year are indicated by an "X". Depending on their response, participants were then presented with a complete list (for each sexual behavior group) and asked to rank the modalities from most likely to least likely to use.

[c]Includes 2 pills within 24 hours before sex and 2 separate 1-pill doses within 2 days after sex.

Participants were also asked to rank their interest in using each of the different potential methods of PrEP administration. The number of modalities that each participant could rank depended on the type of sex he indicated having in the past 12 months. If a respondent said he had both insertive and receptive anal sex in the past year, he could rank all 9 potential modalities. However, if a respondent indicated he only had receptive anal sex in the past 12 months, he was prompted to rank only 7 potential modalities (4 that are applied rectally, 2 applied orally, and an injection).

Statistical Analyses

We did all analyses using SAS v9.4 (SAS Institute Inc). Participants who did not respond to any of the PrEP modality Likert-scale items (n=318) were removed from the analysis dataset. We summarized the likelihood of using each PrEP modality by finding the mean, median, and mode of the 5-point Likert-scale item response. Since these items are ordinal and not interval data, we used nonparametric tests for statistical inferences. We used Wilcoxon signed rank tests to determine whether the reported likelihood of using each of the 8 experimental PrEP modalities differed from the likelihood of using PrEP as a daily oral pill. Each test was considered statistically significant at alpha=.05.

Individual Likert-scale items that asked about topical application sites were summed to create separate Likert scales based on application method and time of application. This resulted in three 10-point Likert scales based on method (penile gel, rectal gel, and rectal suppository) and two 15-point Likert scales based on time of application (before intercourse, after intercourse) [28].

To assess demographic associations with likelihood of using each modality, we dichotomized individual Likert-scale item responses so that "somewhat likely" or "very likely" indicated likelihood of use (versus "somewhat unlikely" or "very unlikely"). Responses of "neither likely or unlikely" were set to missing. We used logistic regression to determine unadjusted odds ratios (ORs) and corresponding confidence intervals for demographic variables of interest.

We used the modified Borda count method to determine the order of preference for the different PrEP modalities [29]. Each modality was assigned a number of points that corresponded to the position in which it was ranked by the participant. The number of points given to a participant's first choice was equal to the number of modalities he actually ranked. We then summed points for each modality to create a collective ranking. Since the number of options presented to an individual depended on the type of sex he reported in the past 12 months (eg, a participant who only reported insertive anal sex was not presented with modalities that are administered rectally), we stratified cumulative rankings by reported sexual practices.

Results

There were 3990 participants who started the online survey and answered eligibility questions. Of 1921 men who met the eligibility requirements for the survey, 4 reported having tested positive for HIV and 493 reported never having oral or anal sex with a man. We further limited the final analysis dataset to the 1106 participants who answered at least one of the questions related to PrEP.

Table 2 summarized the demographic characteristics of the 1106 participants included in the analysis. Almost half of the participants (542/1106, 49.01%) were between 18 and 24 years old and the mean age was 25.2 years. The highest proportion (800/1106, 72.33%) of participants were white, and most had received some level of education past secondary school (1025/1106, 92.68%). The highest proportion of respondents lived in the south (392/1106, 35.44%) but there were at least 200 respondents from all 4 US census regions. The majority of

participants identified as homosexual (965/1106, 87.25%) and reported having had anal sex with a male partner in the past month (964/1106 87.16%). In the previous 12 months, 599 (54.16%) participants reported having both insertive and receptive anal sex with a male partner, 202 (18.26%) reported having only insertive sex, and 163 (14.74%) reported having only receptive sex. There were 138 (12.48%) participants who did not have anal sex in the previous 12 months. The majority of participants (824, 74.50%) had previously heard of people using PrEP to reduce the risk of getting HIV.

Table 2. Demographic and sexual behavior characteristics of 1106 men who have sex with men, aged 18–34 years, participating in an online survey about preexposure prophylaxis (PrEP) for HIV infection, United States, April 2015–July 2015.

Characteristics	n	(%)
Age (years) [a]		
18–24	542	(49.01)
25–29	334	(30.20)
30–34	230	(20.80)
Race/ethnicity		
White, non-Hispanic	800	(72.33)
Black, non-Hispanic	51	(4.61)
Hispanic	181	(16.37)
Other	74	(6.69)
US census region		
Midwest	217	(19.62)
Northeast	251	(22.69)
South	392	(35.44)
West	227	(20.52)
Unknown	19	(1.72)
Highest level of education		
High school or less	80	(7.23)
Some college, associate degree, or technical degree	306	(27.67)
Bachelor degree	384	(34.72)
Any graduate or professional school	335	(30.29)
Unknown	1	(0.09)
Sexual identity		
Homosexual	965	(87.25)
Bisexual	106	(9.58)
Other[b]	35	(3.16)
Type of anal sex with male partner(s) in past 12 months		
Insertive only	202	(18.26)
Receptive only	163	(14.74)
Both insertive and receptive	599	(54.16)
No anal sex	138	(12.48)
Previously heard of PrEP	824	(74.50)

[a]Mean 25.2 years, median 25 years, range 18–34 years.

[b]There were 5 participants who indicated heterosexual, 15 indicated unsure, 14 indicated other, and 1 who did not indicate sexual identity.

Figure 1 displays the stated likelihood of using each PrEP modality to reduce the risk of contracting HIV. Overall, over half of participants stated they would be somewhat likely or very likely to use on-demand pills (685/1105, 61.99%), penile gel (438/755, 58.01% before intercourse and 408/751, 54.33% after intercourse), a periodic injection (575/1091, 52.70%), or a daily oral pill (528/1036, 50.97%). The majority (437/792, 55.18%) of respondents indicated they would be very unlikely

or somewhat unlikely to use a rectal suppository before intercourse.

When the responses were analyzed on a 5-point scale, the mean responses ranged from 2.58 (rectal suppository before intercourse, median 2) to 3.63 (on-demand pills, median 4; Table 3). Compared to the currently available daily pills (median 4), respondents reported a higher likelihood of using on-demand pills (median 4, $P<.001$) or a penile gel before intercourse (median 4, $P=.001$). However, participants reported a significantly lower likelihood of using each of the rectal modalities (all $P<.001$) compared with a daily oral pill.

Figure 1. Stated likelihood of using different preexposure prophylaxis modalities among 1106 men who have sex with men, aged 18-34 years, participating in an online survey, United States, April–July 2015.

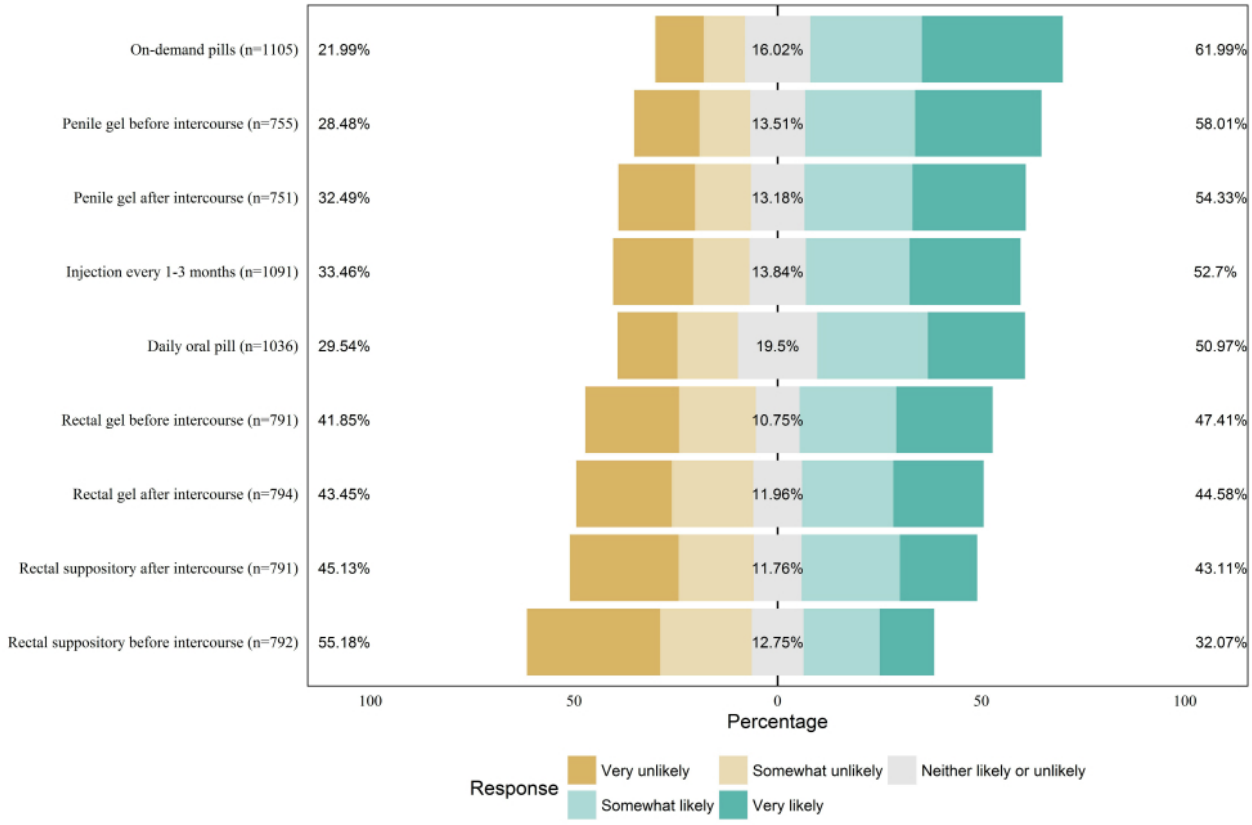


Table 3. Stated likelihood of using different preexposure prophylaxis modalities for HIV infection among 1106 men who have sex with men, aged 18–34 years, participating in an online survey, United States, April–July 2015.

Individual Likert-scale items for each modality[a]	n	Mean (SD)	Median (IQR[b])	Mode	*P*-value[c]
Daily oral pill	1036	3.31 (1.4)	4 (2)	4	Reference
On-demand pills[d]	1105	3.63 (1.4)	4 (2)	5	<.001
Injection[e]	1091	3.27 (1.5)	4 (3)	5	.02
Penile gel before intercourse	755	3.45 (1.4)	4 (3)	5	.001
Penile gel after intercourse	751	3.31 (1.5)	4 (3)	5	.28
Rectal gel before intercourse	791	3.06 (1.5)	3 (2)	4	<.001
Rectal gel after intercourse	794	3.00 (1.5)	3 (2)	1	<.001
Rectal suppository before intercourse	792	2.58 (1.4)	2 (3)	1	<.001
Rectal suppository after intercourse	791	2.90 (1.5)	3 (3)	1	<.001

[a]The 5-point Likert-scale items where 1=very unlikely, 2=somewhat unlikely, 3=neither likely or unlikely, 4=somewhat likely, 5=very likely.

[b]IQR: interquartile range.

[c]*P*-values calculated using Wilcoxon signed rank tests with alpha=.05.

[d]Consisting of 2 pills 24 hours before sex and 2 separate 1-pill doses after.

[e]Every 1–3 months.

On 10-point Likert scales created by combining modalities by topical application methods, the reported likelihood of using a penile gel (median 7) was higher than that of using a rectal gel (median 6, $P<.001$; Table 4). However, the likelihood of using a rectal gel was higher than that of using a rectal suppository (median 6, $P<.001$). There was no statistically significant difference in the reported likelihood of using a topical PrEP modality (gel or suppository) before or after intercourse.

When individual Likert-scale item responses were dichotomized, stated likelihood of using daily oral pills differed by race/ethnicity, age category, and highest level of education (Multimedia Appendix 1). Compared with white participants, black participants had higher odds of reporting a favorable likelihood of using the following modalities: daily oral pills (OR 3.10, 95% CI 1.35–7.13), penis gel before intercourse (OR 6.59, 95% CI 1.54–28.24), penis gel after intercourse (OR 2.82, 95% CI 1.05–7.59), rectal gel before intercourse (OR 4.28, 95% CI 1.73–10.61), rectal gel after intercourse (OR 2.52, 95% CI 1.14–5.58) and a suppository before intercourse (OR 2.69, 95% CI 1.26–5.77).

Figure 2 shows the modified Borda count rankings, stratified by type of anal sex in the past 12 months. On-demand pills were the top-ranked modality for each sexual behavior group. In general, modalities administered orally were ranked highest and modalities administered rectally were ranked lowest.

Table 4. Stated likelihood of using different preexposure prophylaxis topical modalities, by method and time of application, among 1106 men who have sex with men, aged 18–34 years, participating in an online survey, United States, April–July 2015.

Combined Likert scales[a]	n	Mean (SD)	Median (IQR[b])	Mode	*P*-value[c]
Application method [d]					
Penile gel, anytime	750	6.76 (2.7)	7 (4)	10	<.001
Rectal gel, anytime	790	6.06 (2.9)	6 (4)	2	Reference
Rectal suppository, anytime	790	5.48 (2.8)	6 (6)	2	<.001
Time of application [e]					
Before intercourse	585	8.87 (3.8)	9 (6)	3	Reference
After intercourse	584	9.02 (4.0)	9 (6)	3	.14

[a]Created from original 5-point Likert-scale items where 1=very unlikely, 2=somewhat unlikely, 3=neither likely or unlikely, 4=somewhat likely, 5=very likely.

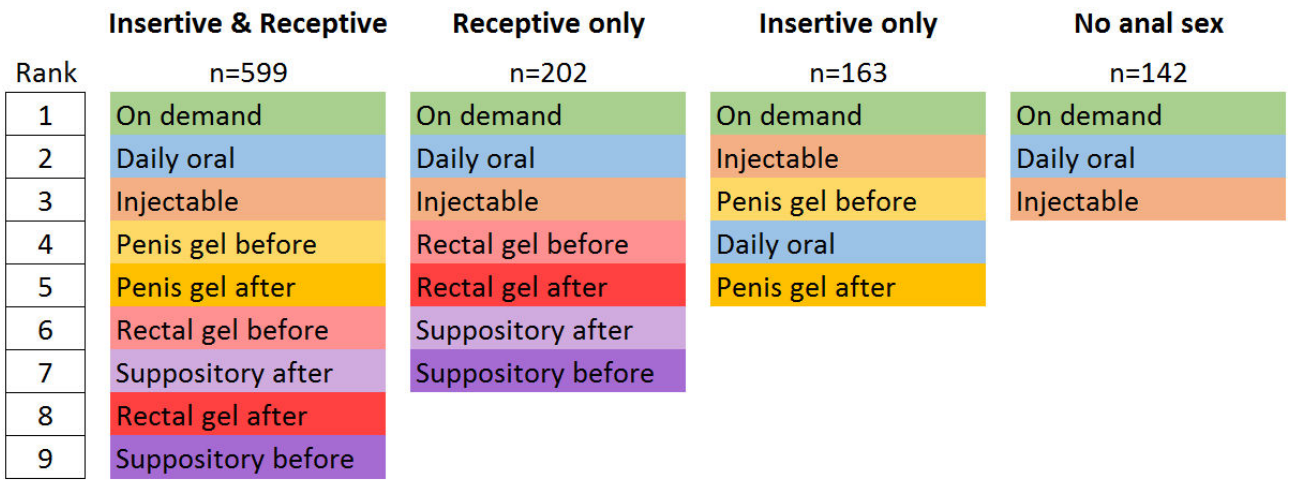
[b]IQR: interquartile range.

[c]*P*-values calculated using Wilcoxon signed rank tests with alpha=.05.

[d]Based on a 10-point scale created by adding the two 5-point Likert-scale items for each application method.

[e]Based on a 15-point scale created by adding the three 5-point Likert-scale items for each time of application.

Figure 2. Modified Borda count ranking of different preexposure prophylaxis modalities, stratified by type of anal sex with a male partner in the past year, among 1106 men who have sex with men, aged 18–34 years, participating in an online survey, United States, April-July 2015. Respondents were presented with a different number of modalities to rank, depending on the type of anal sex they reported in the past 12 months.



Discussion

Our findings indicate that 51% of MSM would be likely to take PrEP as a daily oral pill, which is consistent with the range seen in previous studies in which between 46.1% and 78.5% of MSM said they would be willing to use daily oral PrEP [19,30-33]. The varying range of acceptability is likely a result of some studies stipulating specific scenarios in which PrEP is either offered free of charge, does not cause side effects, or is 80% effective against preventing HIV infection. Some of this variation in acceptability could be related to when the survey was conducted in relation to the public release of the iPrEx results [4]. While our analysis did not investigate reasons against taking PrEP, previous research has indicated that concerns about health (both long-term consequences and immediate side effects), unknown efficacy, possibility of developing drug resistance, cost, and risk perception can all be barriers to use [13,19,20].

These results provide comparative insights on how likely MSM in our sample were to use a variety of hypothetical PrEP modalities. Overall, attitudes toward using PrEP to reduce the risk of contracting HIV were generally neutral or favorable, with a reported likelihood of use ranging from 32% (rectal suppository before intercourse) to 62% (on-demand pills). However, there are some noticeable differences in the likelihoods of using each modality.

The preference for intermittent oral PrEP is evident across this analysis. The highest proportion of participants indicated they would use on-demand pills, which was also the top consensus rank for each of the 4 modified Borda count groups. This preference is particularly of interest when viewed with the growing evidence base demonstrating the efficacy of intermittent (ie, less than daily) oral PrEP [24]. Community surveys have shown that the majority of condomless anal intercourse events appear to be anticipated in advance or infrequent enough to make event-driven or time-driven PrEP regimens feasible [34-36]. Furthermore, Parsons et al recently found that MSM overestimate the likelihood of having sex and are much better at predicting when they would not have sex [37]. This has implications on counseling related to intermittent PrEP use. Parsons et al concluded that counseling messages should encourage individuals to skip a daily dose only when they are sure there is no chance they will have sex the following day. Several forms of slow-acting injectable PrEP are being studied [26], and our study indicates they may have similar acceptability to daily oral pills. For most of the modalities (daily oral pills, both penis gels, both rectal gels, and suppository before intercourse), our study indicated that black participants were more likely than white participants to use them.

It is important to note that attitudes toward topical PrEP modalities differ by administration site. Although gels applied to the penis were generally viewed as acceptable, the 4 rectal modalities were the only Likert-scale items in which more than half of participants reported that they were unlikely to use them. While the combined Likert scales indicated that rectal gels are seen more favorably than rectal suppositories, both were less likely to be used than penile gels. Most research conducted on PrEP in a gel form has focused on vaginal gels for women in Africa [8,22], but there is an ongoing phase 2 trial on rectal gel microbicides in MSM (MTN-017) [38]. Phase 1 research indicated that 75% to 100% of recipients found the experimental gels to be acceptable [39].

Limitations

Several limitations need to be considered when interpreting these results. First, because Likert-scale items are not interval in nature, only the direction of preference can be determined. The magnitude of preference cannot be adequately determined or compared (ie, we cannot say “how much more likely” somebody is to use a single modality over another). Second, we asked participants to state their likelihood of using different modalities without any specifications of cost, efficacy, or possible side effects. As this information (specific to each modality) becomes available, we would expect the reported likelihood of use to change. Furthermore, we asked participants about different modalities of administration based on the type of sex they reported having in the past year. As a result, the attitudes toward any of the modalities administered rectally represent only those of the participants who recently received anal sex (and vice versa for penile application and insertive anal sex). There is potential for selection bias in the analysis dataset. We excluded 318 respondents from the analysis because they did not respond to any of the PrEP modality Likert-scale items. Those excluded respondents were more likely to be African American (8.2% vs 4.6%, P=.009) and more likely to report not having anal sex with a male in the past year (27.8% vs 12.5%, P≤.001), compared with our sample of 1106 participants. Third, our recruitment methods targeted social media users and our sample may not be representative of the general community of MSM.

Since there is not a reference population that can be used as a comparison to assess representativeness, there have not been any other studies that characterized sampling biases using online convenience samples. However, this approach is one of the most common in the field, is consistent with the body of published literature, and allows researchers to quickly collect behavior information on a large number of MSM [40]. Our advertisements targeted young-adult MSM in the United States, which limits the ability to generalize these results to other age groups, regions, or populations affected by the HIV epidemic. Likelihood and acceptability of use is likely to vary depending on perceived risk and cultural acceptability.

Conclusion

Previous research has shown that PrEP has the potential to reduce risk of HIV transmission, but adherence is essential to ensure efficacy. In order to overcome the many barriers to PrEP uptake, we need to develop modalities that are feasible and likely to be used. Among this specific population of MSM in the United States, intermittent systemic approaches seem to be preferred. For topical methods, those that involve application on the penis are preferred over rectal application, and gels are preferred over suppositories. However, further analysis is needed to determine why people would not use particular modalities. Other studies similar to this one need to be carried out for other groups that will be targeted for the newer PrEP modalities.

While there may not be a single PrEP modality that is used by everybody, our study and future ones like it can help determine which technology is most likely to be adopted by specific communities.

Acknowledgments

This work was funded by Eunice Kennedy Shriver National Institute of Child and Health and Human Development (NICHD) Research Grant 1R21-HD074502-01A1. The findings and conclusions in this paper are those of the authors and do not necessarily represent the views of the Centers for Disease Control and Prevention.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Odds ratios for demographic characteristics and stated likelihood of using different preexposure prophylaxis modalities among men who have sex with men participating in an online survey, United States, April-July 2015.

[PDF File (Adobe PDF File), 387KB - jmir_v18i5e111_app1.pdf]

References

1. Beyrer C, Baral SD, van Griensven F, Goodreau SM, Chariyalertsak S, Wirtz AL, et al. Global epidemiology of HIV infection in men who have sex with men. Lancet 2012 Jul 28;380(9839):367-377 [FREE Full text] [doi: 10.1016/S0140-6736(12)60821-6] [Medline: 22819660]
2. Beyrer C, Sullivan P, Sanchez J, Baral SD, Collins C, Wirtz AL, et al. The increase in global HIV epidemics in MSM. AIDS 2013 Nov 13;27(17):2665-2678. [doi: 10.1097/01.aids.0000432449.30239.fe] [Medline: 23842129]
3. HIV Surveillance Report, 2014. Volume 26. Atlanta, GA: Centers for Disease Control Prevention; 2015. URL: http://www.cdc.gov/hiv/pdf/library/reports/surveillance/cdc-hiv-surveillance-report-us.pdf [accessed 2016-03-01] [WebCite Cache ID 6fgoeNpUn]
4. Grant RM, Lama JR, Anderson PL, McMahan V, Liu AY, Vargas L, iPrEx Study Team. Preexposure chemoprophylaxis for HIV prevention in men who have sex with men. N Engl J Med 2010 Dec 30;363(27):2587-2599 [FREE Full text] [doi: 10.1056/NEJMoa1011205] [Medline: 21091279]
5. Baeten JM, Donnell D, Ndase P, Mugo NR, Campbell JD, Wangisi J, Partners PrEP Study Team. Antiretroviral prophylaxis for HIV prevention in heterosexual men and women. N Engl J Med 2012 Aug 2;367(5):399-410 [FREE Full text] [doi: 10.1056/NEJMoa1108524] [Medline: 22784037]
6. Thigpen MC, Kebaabetswe PM, Paxton LA, Smith DK, Rose CE, Segolodi TM, TDF2 Study Group. Antiretroviral preexposure prophylaxis for heterosexual HIV transmission in Botswana. N Engl J Med 2012 Aug 2;367(5):423-434. [doi: 10.1056/NEJMoa1110711] [Medline: 22784038]
7. Choopanya K, Martin M, Suntharasamai P, Sangkum U, Mock PA, Leethochawalit M, Bangkok Tenofovir Study Group. Antiretroviral prophylaxis for HIV infection in injecting drug users in Bangkok, Thailand (the Bangkok Tenofovir Study): a randomised, double-blind, placebo-controlled phase 3 trial. Lancet 2013 Jun 15;381(9883):2083-2090. [doi: 10.1016/S0140-6736(13)61127-7] [Medline: 23769234]
8. Marrazzo JM, Ramjee G, Richardson BA, Gomez K, Mgodi N, Nair G, et al. Tenofovir-based preexposure prophylaxis for HIV infection among African women. N Engl J Med 2015 Feb 5;372(6):509-518 [FREE Full text] [doi: 10.1056/NEJMoa1402269] [Medline: 25651245]
9. Van Damme L, Corneli A, Ahmed K, Agot K, Lombaard J, Kapiga S, FEM-PrEP Study Group. Preexposure prophylaxis for HIV infection among African women. N Engl J Med 2012 Aug 2;367(5):411-422 [FREE Full text] [doi: 10.1056/NEJMoa1202614] [Medline: 22784040]
10. Baeten J, Celum C. Systemic and topical drugs for the prevention of HIV infection: antiretroviral pre-exposure prophylaxis. Annu Rev Med 2013;64:219-232 [FREE Full text] [doi: 10.1146/annurev-med-050911-163701] [Medline: 23020883]
11. van der Straten A, Van Damme L, Haberer JE, Bangsberg DR. Unraveling the divergent results of pre-exposure prophylaxis trials for HIV prevention. AIDS 2012 Apr 24;26(7):F13-F19. [doi: 10.1097/QAD.0b013e3283522272] [Medline: 22333749]
12. Spinner CD, Boesecke C, Zink A, Jessen H, Stellbrink H, Rockstroh JK, et al. HIV pre-exposure prophylaxis (PrEP): a review of current knowledge of oral systemic HIV PrEP in humans. Infection 2016 Apr;44(2):151-158. [doi: 10.1007/s15010-015-0850-2] [Medline: 26471511]
13. Wilton J, Senn H, Sharma M, Tan DH. Pre-exposure prophylaxis for sexually-acquired HIV risk management: a review. HIV AIDS (Auckl) 2015;7:125-136 [FREE Full text] [doi: 10.2147/HIV.S50025] [Medline: 25987851]

14. Koenig LJ, Lyles C, Smith DK. Adherence to antiretroviral medications for HIV pre-exposure prophylaxis: lessons learned from trials and treatment studies. Am J Prev Med 2013 Jan;44(1 Suppl 2):S91-S98. [doi: 10.1016/j.amepre.2012.09.047] [Medline: 23253769]
15. Anderson PL, Glidden DV, Liu A, Buchbinder S, Lama JR, Guanira JV, iPrEx Study Team. Emtricitabine-tenofovir concentrations and pre-exposure prophylaxis efficacy in men who have sex with men. Sci Transl Med 2012 Sep 12;4(151):151ra125 [FREE Full text] [doi: 10.1126/scitranslmed.3004006] [Medline: 22972843]
16. Grant RM, Anderson PL, McMahan V, Liu A, Amico KR, Mehrotra M, iPrEx Study Team. Uptake of pre-exposure prophylaxis, sexual practices, and HIV incidence in men and transgender women who have sex with men: a cohort study. Lancet Infect Dis 2014 Sep;14(9):820-829. [doi: 10.1016/S1473-3099(14)70847-3] [Medline: 25065857]
17. Preexposure Prophylaxis for the Prevention of HIV Infection in the United States-2014: A Clinical Practice Guideline. Atlanta, GA: Centers for Disease Control and Prevention; 2014. URL: http://www.cdc.gov/hiv/pdf/prepguidelines2014.pdf [accessed 2016-03-01] [WebCite Cache ID 6fgoaclcf]
18. Guidance on Pre-exposure Oral Prophylaxis (PrEP) for Serodiscordant Couples, Men and Transgender Women Who Have Sex With Men at High Risk of HIV. Geneva: World Health Organization; 2012. URL: http://apps.who.int/iris/bitstream/10665/75188/1/9789241503884_eng.pdf [accessed 2016-03-01] [WebCite Cache ID 6fgoW8KNu]
19. Golub SA, Gamarel KE, Rendina HJ, Surace A, Lelutiu-Weinberger CL. From efficacy to effectiveness: facilitators and barriers to PrEP acceptability and motivations for adherence among MSM and transgender women in New York City. AIDS Patient Care STDS 2013 Apr;27(4):248-254 [FREE Full text] [doi: 10.1089/apc.2012.0419] [Medline: 23565928]
20. Young I, McDaid L. How acceptable are antiretrovirals for the prevention of sexually transmitted HIV?: A review of research on the acceptability of oral pre-exposure prophylaxis and treatment as prevention. AIDS Behav 2014 Feb;18(2):195-216 [FREE Full text] [doi: 10.1007/s10461-013-0560-7] [Medline: 23897125]
21. Krakower DS, Mayer KH. Pre-exposure prophylaxis to prevent HIV infection: current status, future opportunities and challenges. Drugs 2015 Feb;75(3):243-251 [FREE Full text] [doi: 10.1007/s40265-015-0355-4] [Medline: 25673022]
22. Abdool KQ, Abdool Karim SS, Frohlich JA, Grobler AC, Baxter C, Mansoor LE, CAPRISA 004 Trial Group. Effectiveness and safety of tenofovir gel, an antiretroviral microbicide, for the prevention of HIV infection in women. Science 2010 Sep 3;329(5996):1168-1174 [FREE Full text] [doi: 10.1126/science.1193748] [Medline: 20643915]
23. Kashuba AD, Gengiah TN, Werner L, Yang K, White NR, Karim QA, et al. Genital tenofovir concentrations correlate with protection against HIV infection in the CAPRISA 004 Trial: Importance of Adherence for Microbicide Effectiveness. J Acquir Immune Defic Syndr 2015 Jul 1;69(3):264-269. [doi: 10.1097/QAI.0000000000000607] [Medline: 26181703]
24. Molina J, Capitant C, Spire B, Pialoux G, Cotte L, Charreau I, ANRS IPERGAY Study Group. On-demand preexposure prophylaxis in men at high risk for HIV-1 infection. N Engl J Med 2015 Dec 3;373(23):2237-2246. [doi: 10.1056/NEJMoa1506273] [Medline: 26624850]
25. Castel AD, Magnus M, Greenberg AE. Pre-exposure prophylaxis for human immunodeficiency virus: the past, present, and future. Infect Dis Clin North Am 2014 Dec;28(4):563-583. [doi: 10.1016/j.idc.2014.08.001] [Medline: 25455314]
26. Myers JE, Ellman TM, Westhoff C. Injectable agents for pre-exposure prophylaxis: lessons learned from contraception to inform HIV prevention. Curr Opin HIV AIDS 2015 Jul;10(4):271-277. [doi: 10.1097/COH.0000000000000166] [Medline: 26049953]
27. Sineath R. Methods for improving consent and survey completion in online HIV prevention research with young men who have sex with men. Toronto, ON: JMIR Publications, Inc; 2014 Presented at: Medicine 2.0; 2014; Toronto, ON.
28. Likert R. A technique for the measurement of attitudes. Arch Psychol 1932:55.
29. de Borda JC. Mémoire sur les élections au scrutin.: Mémoires de l'Académie Royale des Sciences; 1781. URL: http://gerardgreco.free.fr/IMG/pdf/MA_c_moire-Borda-1781.pdf [accessed 2016-04-22] [WebCite Cache ID 6gxjCNtvM]
30. Grov C, Whitfield TH, Rendina HJ, Ventuneac A, Parsons JT. Willingness to take PrEP and potential for risk compensation among highly sexually active gay and bisexual men. AIDS Behav 2015 Dec;19(12):2234-2244. [doi: 10.1007/s10461-015-1030-1] [Medline: 25735243]
31. Krakower DS, Mimiaga MJ, Rosenberger JG, Novak DS, Mitty JA, White JM, et al. Limited awareness and low immediate uptake of pre-exposure prophylaxis among men who have sex with men using an Internet social networking site. PLoS One 2012;7(3):e33119 [FREE Full text] [doi: 10.1371/journal.pone.0033119] [Medline: 22470438]
32. Golub SA, Kowalczyk W, Weinberger CL, Parsons JT. Preexposure prophylaxis and predicted condom use among high-risk men who have sex with men. J Acquir Immune Defic Syndr 2010 Aug;54(5):548-555 [FREE Full text] [doi: 10.1097/QAI.0b013e3181e19a54] [Medline: 20512046]
33. Mimiaga MJ, Case P, Johnson CV, Safren SA, Mayer KH. Preexposure antiretroviral prophylaxis attitudes in high-risk Boston area men who report having sex with men: limited knowledge and experience but potential for increased utilization after education. J Acquir Immune Defic Syndr 2009 Jan 1;50(1):77-83 [FREE Full text] [doi: 10.1097/QAI.0b013e31818d5a27] [Medline: 19295337]
34. Volk JE, Liu A, Vittinghoff E, Irvin R, Kroboth E, Krakower D, et al. Sexual frequency and planning among at-risk men who have sex with men in the United States: implications for event-based intermittent pre-exposure prophylaxis. J Acquir Immune Defic Syndr 2012 Sep 1;61(1):112-115 [FREE Full text] [doi: 10.1097/QAI.0b013e31825bd87d] [Medline: 22592590]

35. Stack C, Oldenburg C, Mimiaga M, Elsesser SA, Krakower D, Novak DS, et al. Sexual behavior patterns and PrEP dosing preferences in a large sample of North American men who have sex with men. J Acquir Immune Defic Syndr 2016 Jan 1;71(1):94-101. [doi: 10.1097/QAI.0000000000000816] [Medline: 26371786]
36. van Griensven F, Thienkrua W, Sukwicha W, Wimonsate W, Chaikummao S, Varangrat A, et al. Sex frequency and sex planning among men who have sex with men in Bangkok, Thailand: implications for pre- and post-exposure prophylaxis against HIV infection. J Int AIDS Soc 2010;13:13 [FREE Full text] [doi: 10.1186/1758-2652-13-13] [Medline: 20398261]
37. Parsons JT, Rendina HJ, Grov C, Ventuneac A, Mustanski B. Accuracy of highly sexually active gay and bisexual men's predictions of their daily likelihood of anal sex and its relevance for intermittent event-driven HIV pre-exposure prophylaxis. J Acquir Immune Defic Syndr 2015 Apr 1;68(4):449-455. [doi: 10.1097/QAI.0000000000000507] [Medline: 25559594]
38. Cranston R, Lama J. A Phase 2 Randomized Sequence Open Label Expanded Safety and Acceptability Study of Oral Emtricitabine/Tenofovir Disoproxil Fumarate Tablet and Rectally-Applied Tenofovir Reduced-Glycerin 1% Gel. Pittsburgh, PA: Microbicide Trials Network; 2015 Oct 12. URL: http://www.mtnstopshiv.org/studies/4495 [accessed 2016-03-01] [WebCite Cache ID 6fgoQIyyJ]
39. Mcgowan I, Cranston RD, Duffill K, Siegel A, Engstrom JC, Nikiforov A, et al. A phase 1 randomized, open label, rectal safety, acceptability, pharmacokinetic, and pharmacodynamic study of three formulations of tenofovir 1% gl (the CHARM-01 Study). PLoS One 2015;10(5):e0125363 [FREE Full text] [doi: 10.1371/journal.pone.0125363] [Medline: 25942472]
40. Sanchez T, Smith A, Denson D, Dinenno E, Lansky A. Developing a Web-based HIV behavioral surveillance pilot project among men who have sex with men. Open AIDS J 2012;6:224-231 [FREE Full text] [doi: 10.2174/1874613601206010224] [Medline: 23091580]

Abbreviations

FTC: emtricitabine
HIV: human immunodeficiency virus
MSM: men who have sex with men
OR: odds ratio
PrEP: preexposure prophylaxis
TDF: tenofovir disoproxil fumarate

Edited by G Eysenbach; submitted 01.03.16; peer-reviewed by J Topolovec-Vranic, S LeGrand, C Grov; comments to author 24.03.16; revised version received 29.03.16; accepted 31.03.16; published 19.05.16.

Please cite as:
Hall EW, Heneine W, Sanchez T, Sineath RC, Sullivan P
Preexposure Prophylaxis Modality Preferences Among Men Who Have Sex With Men and Use Social Media in the United States
J Med Internet Res 2016;18(5):e111
URL: http://www.jmir.org/2016/5/e111/
doi:10.2196/jmir.5713
PMID:27199100

Original Paper

The Voice of Chinese Health Consumers: A Text Mining Approach to Web-Based Physician Reviews

Haijing Hao[1], PhD; Kunpeng Zhang[2], PhD

[1]Department of Management Science and Information Systems, University of Massachusetts Boston, Boston, MA, United States

[2]Department of Decision, Operations & Information Technologies, University of Maryland, College Park, College Park, MD, United States

Corresponding Author:
Haijing Hao, PhD
Department of Management Science and Information Systems
University of Massachusetts Boston
100 Morrissey Blvd
Boston, MA, 02125
United States
Phone: 1 8572728162
Fax: 1 6172877887
Email: haohaijing@gmail.com

Abstract

Background: Many Web-based health care platforms allow patients to evaluate physicians by posting open-end textual reviews based on their experiences. These reviews are helpful resources for other patients to choose high-quality doctors, especially in countries like China where no doctor referral systems exist. Analyzing such a large amount of user-generated content to understand the voice of health consumers has attracted much attention from health care providers and health care researchers.

Objective: The aim of this paper is to automatically extract hidden topics from Web-based physician reviews using text-mining techniques to examine what Chinese patients have said about their doctors and whether these topics differ across various specialties. This knowledge will help health care consumers, providers, and researchers better understand this information.

Methods: We conducted two-fold analyses on the data collected from the "Good Doctor Online" platform, the largest online health community in China. First, we explored all reviews from 2006-2014 using descriptive statistics. Second, we applied the well-known topic extraction algorithm Latent Dirichlet Allocation to more than 500,000 textual reviews from over 75,000 Chinese doctors across four major specialty areas to understand what Chinese health consumers said online about their doctor visits.

Results: On the "Good Doctor Online" platform, 112,873 out of 314,624 doctors had been reviewed at least once by April 11, 2014. Among the 772,979 textual reviews, we chose to focus on four major specialty areas that received the most reviews: Internal Medicine, Surgery, Obstetrics/Gynecology and Pediatrics, and Chinese Traditional Medicine. Among the doctors who received reviews from those four medical specialties, two-thirds of them received more than two reviews and in a few extreme cases, some doctors received more than 500 reviews. Across the four major areas, the most popular topics reviewers found were the experience of finding doctors, doctors' technical skills and bedside manner, general appreciation from patients, and description of various symptoms.

Conclusions: To the best of our knowledge, our work is the first study using an automated text-mining approach to analyze a large amount of unstructured textual data of Web-based physician reviews in China. Based on our analysis, we found that Chinese reviewers mainly concentrate on a few popular topics. This is consistent with the goal of Chinese online health platforms and demonstrates the health care focus in China's health care system. Our text-mining approach reveals a new research area on how to use big data to help health care providers, health care administrators, and policy makers hear patient voices, target patient concerns, and improve the quality of care in this age of patient-centered care. Also, on the health care consumer side, our text mining technique helps patients make more informed decisions about which specialists to see without reading thousands of reviews, which is simply not feasible. In addition, our comparison analysis of Web-based physician reviews in China and the United States also indicates some cultural differences.

(J Med Internet Res 2016;18(5):e108) doi:10.2196/jmir.4430

KEYWORDS

online doctor review; physician ratings; text mining; China health consumers

Introduction

Finding information about health care or health care providers through Web-based platforms has been increasing in recent years. According to the 2013 Health Online Report by the Pew Research Center, about 58% of American adults have used the Internet to seek health-related information at least once in the past year. Nearly half (47%) of adults in the United States have searched for their health providers online, 37% have consulted physician-rating sites, and 7% of people who sought information about their health care providers posted one review online [1]. A new study in the United States also found that 59% of survey respondents said that online doctor ratings are "somewhat important" for them, while 19% said they are "very important" when they search for physicians [2]. Similarly, in the Netherlands, about a third of the Dutch population searches for ratings of health care providers [3]. A study of seven European countries showed that, among the people who use the Internet for health-related purposes, more than 40% considered the information provided by these eHealth services to be important when choosing a new doctor [4]. Besides survey studies on the proportion of the population that has used online ratings of health care providers, researchers have studied how people look at or evaluate those online reviews. Research has shown that review style and number of reviews also has an impact on how patients evaluate those online reviews and on patient attitude toward doctors who received reviews online [5].

Health care researchers have examined the phenomena of online doctor ratings quantitatively. One study showed that about 17% of American physicians have been rated on the Internet. Among them, obstetrician/gynecologists were twice as likely to be rated than other specialists [6]. In Germany, 37% of all German physicians were rated on the jameda website in 2012, and most of the rated medical specialties were orthopedists, dermatologists, and gynecologists [7]. In the United Kingdom, 61% of family practice physicians on the National Health Service Choices website were rated, and 69% of ratings showed that patients would recommend their family doctors. Doctors who practice in a larger facility, with a lower proportion of older patients, lower deprivation, higher population density, and who are not in a solo practice are more likely to be rated. Doctors who serve in smaller size facilities but not in a solo practice, with a higher proportion of white patients, lower population density, and patients who are less deprived are more likely to have a higher level of recommendation [8]. In China, about 37% doctors who registered on the "Good Doctor" platform have been reviewed [9]. However, some medical practitioners or health care researchers argue that online reviews might be skewed because the outspoken angry patients are more likely to rate their doctors online. An empirical study showed that physicians who received lower ratings in surveys are less likely to be rated online, but online doctor ratings are positively correlated with patient opinions from surveys and tend to exaggerate at the higher end of the rating spectrum [10]. In China, the majority of quantitative reviews (star ratings) were positive—88% were positive for the doctors' treatment effect measure and 91% were positive for the bedside manner measure [9]. In the United States, most online reviews were quite positive, with an average score 3.93 on a scale of 1-5 [6]. In Germany, two thirds of all ratings are in the best category (very good) [7]. In the United Kingdom, the majority (64%) of the online ratings on their National Health Service Choices website are positive [8].

Besides the quantitative ratings associating with reviews, there are also a large number of online textual reviews about health care providers. They can help both health care providers and researchers understand more about patient opinions about care. Unlike the quantitative scores, textual reviews give patients subjective flexibility and freedom to express opinions on their own experiences and concerns. Analysis of online reviews has already been studied in many other domains. For example, mining product reviews has been quite common and successful in the marketing research or management science area, such as using consumer-generated product reviews to analyze people's online product choices behavior [11], or market structure [12]. However, only a few studies have focused on using text-mining techniques to examine and analyze such largely available textual reviews. One previous US study incorporated latent sentiment analysis into regression analysis and improved state-level health outcome measures [13]. Another study in the United Kingdom applied machine-learning techniques to reviews about hospital service and showed that reviews can be used to predict patient opinions about hospital performance [14]. Based on the authors' knowledge, only one study has employed an automatic text-mining method to capture hidden topics that health care consumers discussed about their health care providers. In particular, they analyzed online doctor reviews in four specialties in New York City: Family/General Practitioner, Dentist, Obstetrics/Gynecology, and Psychiatrist [15]. However, those studies were restricted by the limited number of available reviews.

In this paper, we intend to apply a well-known text-mining method, Latent Dirichlet Allocation (LDA), to examine what Chinese patients said about doctors or health care services by analyzing a large empirical dataset collected from the largest online health community in China, the Good Doctor platform. To the best of our knowledge, there is no empirical study about what Chinese health consumers say about their health care providers online in spite of the fact that China has over half a billion Internet users—the largest population of Internet users in the world [16]—and is already known to have more than one million online reviews of Chinese doctors [17]. This study explores the following research questions: What do Chinese patients say about their doctors online? Do those topics vary across specialty areas? What can health care providers, health care administrators, or policy makers learn from those million reviews? This is particularly important since China's health system has been under reform. In addition, are there any differences in reviews of patient care between China and United States?

Status of Online Doctor Reviews in China

Several online doctor-rating platforms have been created and widely used in China in the past decade. Chinese Medicine Review [18], created in December 2013, focuses on reviewing Chinese traditional medicine doctors. Schedule Web Appointment [19], established in 2010, focuses on online appointment scheduling with doctors across China and is also a platform for reviewing doctors. Among these sites, the "Good Doctor Online" [20] (called "Hao Dai Fu" in Chinese; "Hao" means "good" and "Dai Fu" means "doctor" in Chinese) is the first online doctor review platform in China, initiated in 2006. The Good Doctor is not only the earliest online platform to allow patients to rate and comment on their doctors in many specialty areas, but also the largest one, with more than 300,000 doctors reviewed and one million online reviews [17]. In addition, it provides comprehensive online health-related services such as online appointment scheduling, teleconsultation, patient-doctor forums, and patient clubs for specific doctors.

The Good Doctor was founded with the purpose of helping Chinese health care consumers find "good" doctors for their health-related problems. This is particularly important given that China's health care system has changed substantially, and China has not built any effective referral system since the 1980s. Most people in China have no primary care providers, and Chinese patients usually self-refer to any providers they can afford or they believe to be good [21]. Before the Internet, Chinese consumers either tried their luck to select a doctor randomly or depended on word-of-mouth recommendations. However, many people do not have friends with the same health problems and do not know which doctors or specialist they should see for their health needs. As a result, Chinese consumers face many difficulties while choosing a "good" doctor. In addition, China is short of doctors, as many countries are. The number of physicians per 1000 residents is 1.8 in 2011 for China, 2.5 in the United States, and 2.8 in the United Kingdom [22]. Therefore, in China, it is extremely difficult to get a walk-in visit with a doctor and even harder to see a good specialist in a popular hospital. Many patients or their family members need to go to the hospital very early in the morning to line up because online scheduling or phone appointments are not widely used in China.

The Good Doctor Platform

Since being founded in 2006, the Good Doctor platform has been collecting information about Chinese doctors: demographic information, specialty areas, and technical titles, as well as the associated hospital affiliations, such as name, address, and rank level of the hospital. Technical titles are assigned through an evaluation process under a nationally unified ranking system. It has four levels—from junior to senior—from Resident Physician, Attending Physician, Associate Physician, to Chief Physician. On average, every 5 years a doctor can move one level up in this system. Thus, a title primarily indicates a doctor's work experience and technical skills, which also determines the consultation fee for patients. China's hospital grades are evaluated and determined by a government agency—the National Health Department at the provincial level—and the evaluation standards are based on the hospital facilities, number of beds, technical equipment, quality of care, the doctors' skills, etc [23].

Once a doctor's information is posted on the Good Doctor website, patients can anonymously review the doctors online based on their experiences with those doctors. There are three dimensions on which people can evaluate their doctors on this site: two quantitative measures and one qualitative measure. The two quantitative measures are evaluations of a doctor's treatment outcomes and bedside manners on a 5-level scale, from "Unsatisfied" to "Very Satisfied." The qualitative open-ended textual review can be any description or experience associated with the doctor. In order to control for abusive, inappropriate, or fake reviews, those who leave evaluations online are required to provide phone numbers, seen only by the website administrators, so that the site can confirm the veracity of any questionable reviews.

Methods

Data

We collected 773,279 public reviews from 112,873 doctors on the platform as of April 11, 2014. In total, there were 314,624 doctors from over 3000 hospitals across China on the site. Thus, about 36% of doctors has been rated or commented on by Chinese patients, which has similar rate to that of German doctors' online reviews, 37% [7], but higher than that in the United States, where only 17% of doctors have been rated [6].

After data cleaning, such as removing reviews with inaccurate or incomplete information, we had 731,543 reviews with quantitative ratings, 772,979 reviews with qualitative texts, and 731,264 reviews with both quantitative and qualitative measures. The Good Doctor website includes 9 different major medical specialty areas plus one, called "others," referring to all other less common special areas (see Table 1). We chose the top four specialty areas for analysis in this study: Internal Medicine, Obstetrics (OB)/Gynecology (GYN), Pediatrics, and Chinese Medicine, which received about 23%, 13%, 17%, and 12% of all reviews, respectively. They also have a large number of doctors, with approximately 21%, 19%, 14%, and 11% of doctors, respectively. Table 1 shows that the number of reviews on average that each doctor receives in the specialties of Orthopedics, Oncology, Psychiatry, and Oral Medicine are larger than that in Internal Medicine. But the number of doctors in these areas is much smaller than that in Internal Medicine.

Table 1. Number of reviews and doctors by specialty areas.

Specialty areas	Reviews, n	Reviews, %	Doctors, n	Doctors, %	Reviews per doctor, n
Oncology	7372	0.95	1323	1.12	5.6
Chinese Medicine	90,127	11.66	12,073	10.21	7.5
OB/GYN and Pediatrics	128,762	16.66	16,575	14.01	7.8
Infectious Diseases	3205	0.41	486	0.41	6.6
Internal Medicine	102,441	13.25	22,473	19.00	4.6
Orthopedic	3865	0.50	498	0.42	7.8
Others	240,099	31.06	36,281	30.67	6.6
Psychiatry	6429	0.83	1056	0.89	6.1
Oral Medicine	16,346	2.11	2679	2.26	6.1
Surgeon	174,302	22.55	24,846	21.00	7.0
Total	772,948	100	118,290	100	6.5

Topic Modeling

Topic modeling is a sophisticated text-mining technique appropriate for our research task, which is understanding the voice of online Chinese health care consumers by identifying topics on the Good Doctor platform. Topic modeling is a statistical method to uncover abstract topics from a collection of documents [24]. For example, if a document includes flu as a topic, this document is likely to contain related words such as “cold,” “fever,” “cough,” “sneezing”, etc. If a document is about a topic of surgery, then “pain,” “operation,” “surgeon,” “incision,” etc, would co-appear often with high probabilities. Note that the name of the topic is abstracted and summarized by researchers (such as the topics “flu” or “surgery”) based on the most frequently appearing keywords because computer algorithms can find only the pattern of which keywords cluster statistically but cannot summarize what topic those keywords represent. Also, a document usually has a mixture of different topics. Topic modeling can capture those topics in a statistical way by using different algorithms. We used LDA to analyze Chinese consumers’ reviews about their health care providers. LDA has been widely used in various domains, including Web-mining [25], video analysis [26, 27], spam filtering [28], and natural language processing [15, 24, 29, 30]. It is a generative probabilistic model and was first presented for topic discovery by [31], as shown in Figure 1.

β is the parameter of the Dirichlet prior on the per-topic word distribution. α is the parameter of the Dirichlet prior on the per-document topic distributions. $\theta^{(d)}$is the topic distribution for document d(eg, a review), and $z$is the topic assignment for word $w$in a document. $\varphi^{(z)}$is the word distribution for topic z. $w$is the word, while $D$is the number of documents. $N_d$is the number of words in a document d, and $T$is the number of topics. The LDA model assumes the following generative process for a document d=(w_1, . . . w_{nd},) containing $N_d$words from a vocabulary consisting of Vdifferent terms, $w_i$is the i^{th}word for all i=1, . . . , N_d. It consists of the following three steps:

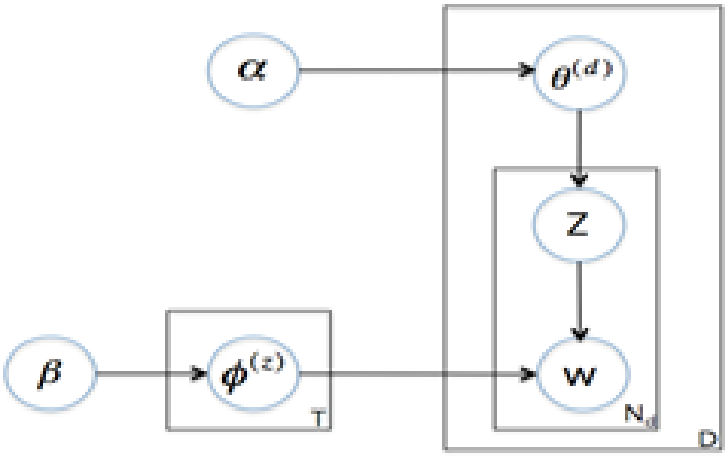
1. The proportions $\varphi$of the word distribution for the topic $z$is determined by $\varphi^{(z)}$~ Dirichlet(β).
2. The proportions $\theta$of the topic distribution for the document $d$are determined by $\theta^{(d)}$~ Dirichlet(α).
3. For each of the $N_d$words: (a) Choose a topic z~ multinomial($\theta^{(d)}$), and (b) Choose a word $w_i$from a multinomial probability distribution conditioned on the topic z: $P(w_i | z, \varphi^{(z)})$

The number of topics in LDA has to be fixed a priori. In this paper, we split the dataset into two parts: 90% training dataset and 10% testing dataset. The optimal number of topics is determined by the perplexity of the trained model on testing dataset. The key inferential problem that we need to solve in order to use LDA is that of computing a posterior distribution of the hidden variables given in a document: $P(\theta, \varphi, z \mid w, \alpha, \beta) = P(\theta, \varphi, z, w \mid \alpha, \beta)/ P(w \mid \alpha, \beta)$.Unfortunately, this distribution is difficult to compute [32]. Although the posterior distribution may not be possible for an exact inference, a wide variety of approximate inference algorithms can be considered for LDA, including varying approximation, Gibbs sampling, and expectation propagation.

The Chinese language is very different from the English language: there are about 3000 basic and commonly used Chinese characters. One single Chinese character usually cannot convey a complete and accurate meaning. At least two or three Chinese characters combined are needed. Therefore, in this study, we need to do some preprocessing before directly applying LDA. We first employ a Chinese character segmentation algorithm implemented by LingPipe [33] to extract meaningful tokens, including removing non-ASCII characters, non-Chinese characters, etc. The extracted tokens may have various lengths from one to a possibly very large number. Each token is considered an atomic entity, meaning that all characters in each token will not be separated for further processing. Then we remove nonsense words, such as stop words in Chinese (eg, of, I, we), and many highly frequent words (eg, doctor, physician, hospital). Finally, we filter out tokens with only one single Chinese character (not meaningful) or those with more than four Chinese characters (likely containing more than one meaning). After this data cleansing, we conduct the following two analyses for each specialty: (1) count the frequency for each

unique token, and (2) run the LDA algorithm on reviews for all doctors in each of the four areas to find the top 10 topics, each of which is represented using 10 words with the highest probability within that topic.

Figure 1. LDA model for topic discovery.



Results

Descriptive Statistics

Figure 2 shows the distribution of the percentage of doctors over the number of reviews a doctor received for all four specialty areas. We found that about 36-45% of doctors received one review, 36-39% of doctors received 2- 5 reviews, 8-12% received 6-10 reviews, and 16-27% of doctors received 6 or more reviews. In extreme cases, several doctors received more than 500 reviews. From Figure 2, we also find that the distribution patterns are similar for all four specialty areas.

Figure 3 shows the distribution of the review volume from 2007-2013. We did not include reviews for 2006 and 2014 because data are incomplete for the entire calendar year. Figure 3 shows that the number of reviews has been relatively increasing. The trend of the number of reviews per doctor is similar for all four specialty areas over the years as shown in Figure 4.

Table 2 shows the descriptive statistics of review length. If we use regular stoppers to split reviews, such as periods, exclamation marks, or question marks, each review has 3-4 sentences on average for all four specialty areas. Compared to American patients' average description length, which is about 4 sentences [15], Chinese patients' reviews seem to be slightly shorter. But, the sentence structure of Chinese is different from English. For example, one Chinese sentence can run very long involving one or more topics separated by commas, which is rare in English. Therefore, to accurately understand the review length, we examine the number of Chinese characters instead. The average number of Chinese characters a review contains is between 85 and 102, which is equivalent to about 40 English words. Based on the authors' published translated books, the translation rate between Chinese and English is usually 2 to 1, that is, a sentence of 20 Chinese characters can be translated into a sentence of 10 English words. Or if you randomly select a Chinese sentence and put it into Google translation, the translation rate is similar. Table 2 shows that the median length of a review is about 60-70 characters. This indicates more than 50% of the reviews are longer than 30 English words. For some extreme cases, it can be up to over 1700 words for Internal Medicine, OB/GYN, and Pediatrics.

Table 2. The descriptive statistics of review length.

	By number of regular stoppers			By number of Chinese characters		
	Average length of reviews	Median length of reviews	Maximum length of reviews	Average length of reviews	Median length of reviews	Maximum length of reviews
Internal medicine	3.3	3	96	85	61	1766
Surgery	3.5	3	81	95	66	1030
OB/GYN and pediatrics	3.6	3	112	101	72	1730
Chinese medicine	3.7	3	77	102	74	1188

Figure 2. Distribution of the number of reviews a doctor received.

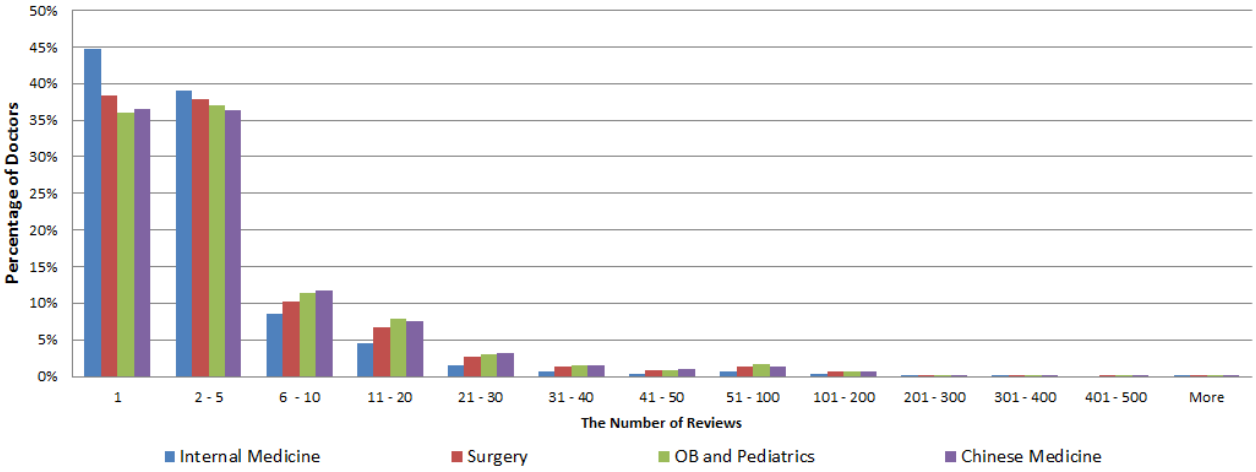


Figure 3. Total number of reviews by specialty over time.

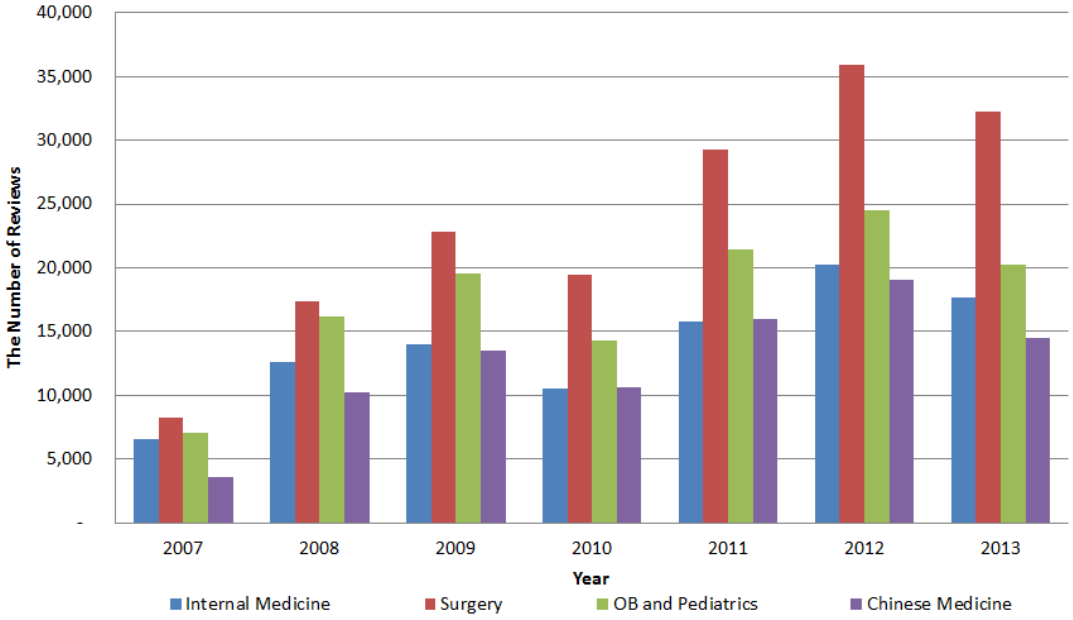
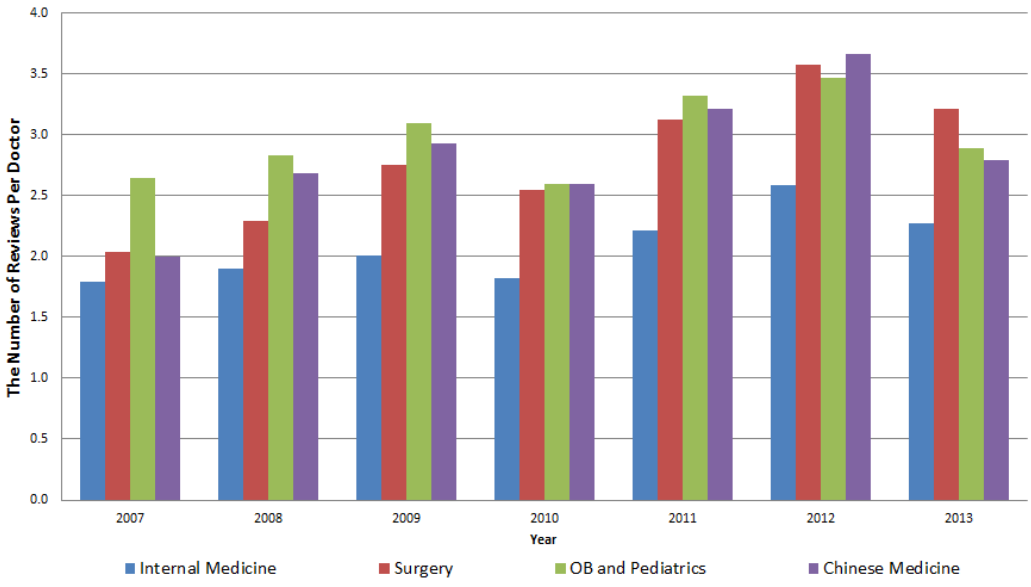


Figure 4. Number of reviews per doctor by specialty over time.



Topic Modeling Result

Based on the LDA algorithm, we identified the top 10 popular topics for each of the four selected specialty areas. Due to space limitations, we present the top 5 popular topics and translate them into English. There are several overlapping or similar topics across specialty areas. The title of each topic is summarized by authors based on the set of keywords returned by the LDA algorithm. For each topic, we show only 10 key words having higher probabilities under that topic. In Tables 3-6, we present the topic summary, the corresponding English translation of the top 10 keywords, and one example review for each topic for the selected four specialty areas, respectively. Based on the assumption of LDA model, each review is a mixture of topics with different probabilities. We select one example review under a topic with the maximum probability.

Table 3. Internal medicine.

Topic	Examples
Topic 1: Treatment effects	
Keywords	Now, effect, always , already, better, after, significantly, one time, one year, many years
Example	I took my mother to see a doctor, and she (mom) had this problem for over 10 years but we never found what caused this problem. Fortunately, we met a student who knew him (doctor) and we visited him for his reputation, diagnosed the problem, and got very good results after.
Topic 2: Technical skills	
Keywords	Medical skill, professional integrity, manner, noble, excellent, super, technique, great, service, experience
Example	Dr. Liu Shengyun is a good doctor, his consultation process was very thorough, and he has very nice manner, excellent technical skills, and great professional integrity.
Topic 3: Story of registration	
Keywords	Time, out-patient, expert, online, because, saw, consulting, today, registration, visit
Example	I registered for my sister. It was extreme difficult. I almost died in the crowd because too many people were in the registration window. Luckily I got a registration ticket. My sister doesn't live in this city. If she comes to register by herself, I guess that she will never get the registration ticket. The medication works well. But she has a swollen face after the medication. At the beginning the doctor said that we should do a follow-up visit in 2 months, so tomorrow morning I will come to fight with the crowd again.
Topic 4: Bedside manner	
Keywords	Patience, manner, serious, responsible, careful, problem, cautious, particularly, every time, warm
Example	Dr. Liu not only has super great technical skills but is very warm and patient with his patients, as if treating his own friends! He is the best doctor that I ever met.
Topic 5: Story of finding doctors	
Keywords	One, this year, found, introduce, child, Beijing, later, Xiehe hospital, last year, start
Example	My friend introduced this doctor (to me) and said that this doctor is good for rheumatism. He diagnosed me after reviewing (my) medical record materials and said that only medication may not be ideal, then introduced an intervention doctor, but there were too many patients there, I didn't have time to ask more questions. Hope next time I will have time to ask.

Table 4. Surgery.

Topic	Examples
Topic 1: Bedside manner	
Keywords	Patient, symptom, serious, manner, problem, situation, cautious, very careful, responsible, careful
Example	Chief Physician Li is very kind, very careful and responsible, and he treats his patients like his own family, very patient, cautious, and he is always very kind, examining carefully, answering questions seriously, and my surgery was very successful. Chief Physician Li is a highly skilled doctor and high professional integrity doctor.
Topic 2: Appreciate the surgery results	
Keywords	Surgery, success, father, mother, removal, check in hospital, surgical
Example	My father had surgery three days ago, before and after the surgery, Chief Physician Jiao Wenjie with his team gave patients the most comfortable and careful care. As the whole family of the patient, we saw all of that care and wanted to say thanks from the bottom of our heart!
Topic 3: General appreciation	
Keywords	Appreciate, thanks, hello, mom, whole family, hello, hope, health, child, here
Example	Your amazing hands are a miracle, your kind smiling is as warm as Spring, you are the good luck star for patients, really appreciate you, Dr. Liao Jianna. Best wishes: Good people have a good life.
Topic 4: Description of symptoms	
Keywords	Because, self, know, feel, when, but, always, this, what, many
Example	I had a bad heart and visited Dr. Li and did some exams, then I knew that I had some problems with aorta valve and atrial fibrillation, then I stayed at hospital, took medication for my heart, until I reached certain conditions, I had surgery.
Topic 5: Story of finding doctor	
Keywords	Time, outpatient, online, found, introduce, expert, saw, friend, shanghai, through
Example	I have visited all the outpatients that have departments of cerebral surgery in Qingdao, but later a friend introduced Chief Physician Jiao to me, after Chief Physician Jiao's patient explanation, which helped us laypersons understand the symptoms and the best treatment. Because our budget is tight, Chief Physician Jiao designed a particular operation plan for me to ensure the best surgery outcomes and least cost. This touched us deeply. Chief Physician Jiao is a reliable and respectful good doctor.

Table 5. Gynecology/OB and pediatrics.

Topic	Examples
Topic 1: General appreciation	
Keywords	Appreciate, thanks, hello, hope, healthy, health, this, whole family, mom, smooth
Example	I really appreciate Chief Physician Wang, who cured my child's decreasing white cells. Nice manner, right medication, cost less. Really appreciate. Wish good health to Chief Physician Wang.
Topic 2: Story of registration	
Keywords	Time, online, expert, because, outpatient, saw, today, visit, therefore, registration
Example	It is very difficult to get through registration, and only one day out of a week which has 10 registration tickets, and (I) need to line up at 5 o'clock for the registration, and finally get the 18th ticket. One day only has 20 tickets. The doctor is very patient and examines very me carefully, and on average (he can) talk 15 minutes with one patient. Also, the medication is not expensive. It is said that he is an expert of this field, and (I) hope a good doctor like this can have more outpatient time. (I) Trust his technical skills.
Topic 3: Story of treatment	
Keywords	Exam, result, know, what, when, this, what, one time, others
Example	My child suddenly got laryngotracheal bronchitis, stayed in a hospital for a week, and kept coughing. Every time my child coughs, doctors would say we need to check for asthma. This time when I took my child to get (medical) exam, Dr. Bo was very kind, and also said very definitely that this is not asthma, no need to do other exams, but only a little medicine needed and (ask the child) take it before bed (the child is difficult with taking any medicine). Only two medicines, and each is a half pill, my child was cooperative for taking (this) medicine (in the past, there were several medicines which my child hated), the effect is good, and I really appreciate it! Excellent technical skill and great professional integrity!
Topic 4: Story of finding doctor	
Keywords	Always, pregnant, one time, introduce, found, many, later, period, friend, because
Example	Because a friend at Provincial Hospital's introduced me, I visited Dr. Feng. I have an ovarian cyst on the left side and pelvis fluid, so I cannot get pregnant. After one month of medication, the ultrasound shows that both problems are solved. Dr. Feng is a doctor with a parent's heart. But many times she was too busy to check patients. Later I had 3 months of Gestrinone Capsules. Ultrasound shows both ovarian cyst and pelvis fluid disappeared. But the lower left side of my belly still has a little pain, and when I touched the left side, I could feel something there, it is different from the right side, but the doctor was too busy and only looked at the ultrasound result, did not check. Because of work, I moved to Shanghai, and now over half year, I felt menstrual cramps last month and Shanghai People's hospital examined and diagnosed it as ovarian cyst. Now after the whole consultation and treatment time, I can say I am exhausted. Doctors at Shanghai said no medication for menstrual cramps. I want to go back to Guangdong for Dr. Feng, her work attitude and technical skills are good, as everyone can see.
Topic 5: Story of surgery	
Keywords	Surgery, uterus, in-patient, year and month, recover, success, fibroids, leave hospital, follow up, found out
Example	Hello, I am one of your many patients, since I got Uterine fibroids, I have seen many doctors and they all said that I should remove my uterus, I received the recommendation from an acquaintance. I come for the reputation, and Dr. Du of Zhejiang People Hospital is excellent.

Table 6. Chinese medicine.

Topic	Examples
Topic 1: Bedside manner	
Keywords	Patience, symptom, manner, serious, cautious, problem, careful, situation, query, every time
Example	This doctor is a very lovely doctor, my friend accompanied me to see him, once (my friend) saw that this doctor is on duty, and (my friend) got a registration ticket immediately, and said he would be very patient to explain everything.
Topic 2: Technical skills	
Keywords	Manner, technical skill, professional integrity, also, noble, excellent, this, responsible, super, great
Example	Dr. Ni is a good doctor for his true values. He has a very kind manner to his patients, very nice, very considerate to his patient, excellent technical skills, rich experience! I have been to many hospitals, have seen many doctors, but a doctor like Dr. Ni with such noble professional integrity is hard to find. You can just go and see him, once you see him, you will know what a good doctor means!
Topic 3: Description of symptoms	
Keywords	Many years, myself, serious, because, pain, symptom, mother, cannot, acupuncture, go through
Example	Appreciate Dr. Sun cured my mom's arthritis of the shoulder which she had for many years. Also gave me a very targeted treatment. Now, I come to see Dr. Sun again for treatment.
Topic 4: Story of finding doctors	
Keywords	Chinese medicine, online, found, Chinese medicine hospital, year month, at that time, saw, this year, Beijing, found out
Example	I work in Shanghai and always have stomach problems. Doctors in Shanghai said that is "Reflux Esophagitis" and I have had both western medication and Chinese traditional medication from Level 3 A hospital in Shanghai for almost 5 years, but it didn't cure it and I had to keep taking medication, which really confused me. I found President Ji of Nantong Chinese Medicine Hospital, so I went to Star Doctor Department at Nantong Chinese Medicine Hospital to see him. President Ji did Gastroscopy for me and found it was "superficial gastritis". Now I have used his western medication and Chinese medication for about 2 months, I am much better. I am very happy now.
Topic 5: Concern about child's health	
Keywords	Child, cough, start, son, daughter, cold, baby, later, every time, kid
Example	Hello, Dr. Hu! Today I will write a belated appreciation letter. Twenty years ago I took my son to see you for his respiratory system asthma, at that time my son was only 2-3 years old, and he got colds frequently, there was a special sound when he coughed, sometimes, and he would get coughing problems 2-3 times in a month. But after seeing you, (after) about 3-4 times in total, (Chinese medicine and acupuncture). Then this problem never came back, sometimes (he) may get cold but no coughing problem any more. Our entire family really appreciates that you cured my son's problem. Thanks! I always remember your name, if my colleagues or friends' children have this problem, I will introduce you to them. You are a really high skilled doctor.

Table 7. Topics comparison across specialty areas (X means reviews of doctors under a corresponding specialty area largely describe that topic).

	Internal medicine	Surgery	OB-Pediatrics	Chinese Medicine
Treatment effects	X			
Technical skills	X			X
Appreciate the surgery result		X		
Story of treatment			X	
Story of surgery			X	
Bedside manner	X	X		X
Story of registration	X		X	
Story of finding doctors	X	X	X	X
General appreciation		X	X	
Description of symptoms		X		X
Concern about children's health				X

Discussion

Principal Findings

We found some common and distinct topics among the four specialty areas. For example, in Table 7, we can see that the most common topic across four specialty areas is the "story of finding doctors," which is not a surprise given the following. First, the goal of the Good Doctor platform is to help Chinese patients find good doctors or good specialists for their health problems. Describing how to find good doctors in reviews on this platform should be common. Second, we know that there is no mature primary care systems or professional referral systems in China. This may cause unexpected difficulty for a Chinese patient to figure out which specialist they should see for their medical concerns. Finally, due to the shortage of doctors, 1.8 doctors per 1000 people in China (compared with 2.5 in the United States and 2.8 in the United Kingdom [22]), obtaining a "ticket" for the registration system to see a doctor is always challenging, which results in many complaints. Some of the randomly selected examples in Tables 3-6 also show how difficult it is to see a doctor or how busy a doctor is.

Our findings also show that some topics are quite common and are included across specialty areas, for example, "technical skills" and "bedside manner." This is not only because they are a focus of patient care, but also that the platform elicits such kinds of reviews. The Good Doctor platform asks reviewers to give rating scores based on these two dimensions before writing text reviews. "General appreciation" and "description of symptoms" are another two common topics across specialty areas. All other topics in the table are found only within one specific specialty. For example, "treatment effects" is seen more in Internal Medicine. Reviews of doctors in Surgery focus on "appreciate the surgery results." "Concern about children's health" is reflected more by reviews in the specialty of Chinese Medicine, and this may suggest that Chinese parents prefer to take their children to see Chinese medicine specialists to avoid the potential side effects from western medicine.

We also conducted a comparison between Chinese doctor reviews and American doctor reviews. Topics extracted from both are sometimes different but quite close for similar medical specialties. For example, reviews under Family/General Practitioner in the United States were related more to topics like "manner" and "competence," while Chinese patients paid more attention to "bedside manner," "technical skills," and "treatment effects." For the specialty area of OB in the United States (OB/GYN and Pediatrics in China), we found that American patients talked more about "manner," "anecdotal," "attention," and "recommendation," while Chinese patients focused more on various topics, such as "stories on treatment," "surgery," "finding doctors," and "general appreciation." We also found that many reviewers in the United States recommended doctors explicitly if they were satisfied with their experiences. This may indicate that they consciously realize that other patients may read their posts later. For Chinese patients, they use many polite words to show their appreciation to their doctors directly by addressing the doctor and some also explicitly display their own names and phone numbers, which may indicate that they wish their doctors to recognize them through reviews to receive better treatment next time. Those differences may result from cultural differences. In addition, "attention" is commonly seen in the American doctor reviews. But in this study, we did not specifically have such a topic. It may be included in "bedside manner" and "general appreciation." Finally, American patients specifically discussed "cost" under the specialty of dentist, and "schedule" for psychiatrist [15]. We did not include these two specialty areas because first, the current Good Doctor platform does not have a separate category called dentist. It is included in the category of oral medicine. Second, the total number of reviews on psychiatrists is too small, with only about 6000 reviews across 1000 doctors.

Limitations

There are limitations in this study. First, LDA has been used to extract hidden topics [34]. LDA is mainly based on the frequency of co-occurrence of words under similar topics. It might not able to identify some topics that are mentioned by very few reviewers, for example, some emerging topics. Second, our data were collected only from the Good Doctor platform, which might lead to some limitations in data source setting and our methodology. However, the Good Doctor is the largest primary platform of Web-based doctor reviews in China. Also, we want to note that when patients post their reviews on the Good Doctor platform, they are asked to leave their phone number to the webmaster in case any questionable comments need to be verified. This strategy is to prevent any dishonest comments or automatic robotic work. However, some patients might be a little hesitant to make negative comments because of this feature.

Conclusions

To summarize, Web-based physician review platforms are a good channel for Chinese patients to express their opinions and share their experiences. Topics extracted from those user-generated reviews can provide more understanding of what patients posted. It can also help health care policy makers and health care providers monitor and adjust their policies or resources to better serve their people and improve the quality of health care.

Our study makes several contributions. First, to the best of our knowledge, this is the first text-mining study to understand the voice of Chinese health care consumers by analyzing a large number of Web-based physician reviews. Data were collected from the largest online health care platform in China. Second, we discovered topics from over 500,000 online textual reviews and compared them across specialty areas. Automatic topic analysis provides patients a way to know more about doctors in order to help them make decisions on which specialist they should see. It also helps health care providers or health care policy makers understand patients' concern or complaints, thus they can adjust their policies or resources to better serve people and improve the quality of health care. Given the fact that the number of online reviews is dramatically increasing, it is simply not practical for a person to read and analyze all the reviews that have been posted. We also compared topics extracted from Chinese doctor reviews and American doctor reviews. Finally,

our study empirically demonstrates that Chinese consumers care more about registration, doctors' bedside manner, and technical skills along with other topics.

There are a few potential studies that could stem from this research. First, we can extend this topic modeling study to Web-based physician reviews for all specialty areas to provide health care providers and researchers with better insight into consumers' thoughts regarding different medical specialties. Second, we can apply this text-mining technique to similar reviews in other countries to determine whether there are any differences across cultures. Third, it would be interesting to further investigate the relationship between Web-based physician reviews and the quality of care provided by health care professionals.

Acknowledgments

We wish to express our gratitude for the continuing support of Mr. Hang Wang, the founder of the Good Doctor platform, and his technical team for this research.

Conflicts of Interest

None declared.

References

1. Kadry B, Chu LF, Kadry B, Gammas D, Macario A. Analysis of 4999 online physician ratings indicates that most patients give physicians a favorable rating. J Med Internet Res 2011;13(4):e95. [doi: 10.2196/jmir.1960] [Medline: 22088924]
2. Hanauer DA, Zheng K, Singer DC, Gebremariam A, Davis MM. Public awareness, perception, and use of online physician rating sites. JAMA 2014 Feb 19;311(7):734-735. [doi: 10.1001/jama.2013.283194] [Medline: 24549555]
3. Van de Belt TH, Engelen LJ, Berben SA, Teerenstra S, Samsom M, Schoonhoven L. Internet and social media for health-related information and communication in health care: preferences of the Dutch general population. J Med Internet Res 2013;15(10):e220 [FREE Full text] [doi: 10.2196/jmir.2607] [Medline: 24088272]
4. Santana S, Lausen B, Bujnowska-Fedak M, Chronaki C, Kummervold PE, Rasmussen J, et al. Online communication between doctors and patients in Europe: status and perspectives. J Med Internet Res 2010;12(2):e20 [FREE Full text] [doi: 10.2196/jmir.1281] [Medline: 20551011]
5. Grabner-Kräuter S, Waiguny MKJ. Insights into the impact of online physician reviews on patients' decision making: randomized experiment. J Med Internet Res 2015;17(4):e93 [FREE Full text] [doi: 10.2196/jmir.3991] [Medline: 25862516]
6. Gao GG, McCullough JS, Agarwal R, Jha AK. A changing landscape of physician quality reporting: analysis of patients' online ratings of their physicians over a 5-year period. J Med Internet Res 2012;14(1):e38 [FREE Full text] [doi: 10.2196/jmir.2003] [Medline: 22366336]
7. Emmert M, Meier F. An analysis of online evaluations on a physician rating website: evidence from a German public reporting instrument. J Med Internet Res 2013;15(8):e157 [FREE Full text] [doi: 10.2196/jmir.2655] [Medline: 23919987]
8. Greaves F, Pape UJ, Lee H, Smith DM, Darzi A, Majeed A, et al. Patients' ratings of family physician practices on the internet: usage and associations with conventional measures of quality in the English National Health Service. J Med Internet Res 2012;14(5):e146 [FREE Full text] [doi: 10.2196/jmir.2280] [Medline: 23076301]
9. Hao H. The development of online doctor reviews in China: an analysis of the largest online doctor review website in China. J Med Internet Res 2015;17(6):e134 [FREE Full text] [doi: 10.2196/jmir.4365] [Medline: 26032933]
10. Gao G, Greenwood B, Agarwal R, McCullough J. Vocal Minority and Silent Majority: How Do Online Ratings Reflect Population Perceptions of Quality? MIS Quarterly 2015 2015;39(3):565-589.
11. Archak N, Ghose A, Ipeirotis P. Deriving the Pricing Power of Product Features by Mining Consumer Reviews. Management Science 2011;57(8):1485-1509.
12. Netzer O, Feldman R, Goldenberg J, Fresko M. Mine Your Own Business: Market-Structure Surveillance Through Text Mining. Marketing Science 2012;31(3):521-543.
13. Wallace B, Paul MJ, Sarkar U, Trikalinos T, Dredze M. A large-scale quantitative analysis of latent factors and sentiment in online doctor reviews. J Am Med Inform Assoc 2014;21(6):1098-1103 [FREE Full text] [doi: 10.1136/amiajnl-2014-002711] [Medline: 24918109]
14. Greaves F, Ramirez-Cano D, Millett C, Darzi A, Donaldson L. Use of sentiment analysis for capturing patient experience from free-text comments posted online. J Med Internet Res 2013;15(11):e239 [FREE Full text] [doi: 10.2196/jmir.2721] [Medline: 24184993]
15. Brody S, Elhadad N. Detecting Salient Aspects in Online Reviews of Health Providers. 2010 Presented at: AMIA 2010 Symposium Proceedings; Nov. 13-17, 2010; Washington, DC p. 202.
16. The World Bank. URL: http://www.webcitation.org/6WWUrs3OY [WebCite Cache ID 6WWUrs3OY]
17. The Economist website. URL: http://www.webcitation.org/6WWV4pi2N [WebCite Cache ID 6WWV4pi2N]
18. Chinese Medicine Review. URL: http://www.zydp.org/ [accessed 2016-04-26] [WebCite Cache ID 6h4CkhNMZ]
19. Schedule Web Appointment. URL: http://www.guahao.com/ [accessed 2016-04-26] [WebCite Cache ID 6h4CwAvRA]

20. The Good Doctor. URL: http://www.haodf.com/ [accessed 2016-04-26] [WebCite Cache ID 6h4DARmey]
21. Eggleston K, Ling L, Qingyue M, Lindelow M, Wagstaff A. Health service delivery in China: a literature review. Health Econ 2008 Feb;17(2):149-165. [doi: 10.1002/hec.1306] [Medline: 17880024]
22. World Health Organization's Global Health Workforce Statistics. 2011. URL: http://www.webcitation.org/6gutU5Jna [WebCite Cache ID 6gutU5Jna]
23. National Health and Family Planning Commission. Hospital Grade Level Management Evaluation. Beijing, China: National Health and Family Planning Commission; 1998.
24. Brody S, Elhadad N. An unsupervised aspect-sentiment model for online reviews. 2010 Presented at: Conference of the North American Chapter of the Association for Computational Linguistics (NAACL-HLT); June 1-6, 2010; Los Angeles, CA p. 804-812.
25. Mei Q, Liu C, Su H, Zhai C. A Probabilistic Approach to Spatiotemporal Theme Pattern Mining on Weblogs. 2006 Presented at: 15th International World Wide Web Conference (WWW'06); May 22-26, 2006; Edinburgh, Scotland, UK.
26. Niebles J, Wang J, Li F. Unsupervised Learning of Human Action Categories Using Spatialtemporal Words. In: Proc BMVC. 2006 Presented at: British Machine Vision Association (BMVC); Sept. 4-7, 2006; Edinburgh, UK. [doi: 10.5244/C.20.127]
27. Wang Y, Sabzmeydani P, Mori G. Unsupervised Activity Perception by Hierarchical Bayesian Model. In: Proc CVPR. 2007 Presented at: Computer Vision and Pattern Recognition; June 18-23, 2007; Minneapolis, Minnesota.
28. Biro I, Siklosi D, Szabo J, Benczur A. Linked Latent Dirichlet Allocation in Web Spam Filtering. 2009 Presented at: 5th International Workshop on Adversarial Information Retrieval on the Web (AIRWeb); Apr. 21, 2009; Madrid, Spain.
29. Boyd-Graber J, Blei D. Syntactic Topic Models. In: NIPS. 2009 Presented at: Neural Information Processing Systems; Dec. 7, 2009; Vancouver, BC.
30. Blei D, Griffiths T, Steyvers M, Tenenbaum J. Integrating Topics and Syntax. In: Proc. Neural Information Processing Systems. 2005 Presented at: Neural Information Processing Systems; Dec. 7-12, 2005; Montreal, Canada.
31. Blei D, Ng A, Jordan M. Latent Dirichlet Allocation. Journal of Machine Learning Research 2003:993-1022.
32. Dickey J. Multiple hypergeometric functions: Probabilistic interpretations and statistical uses. Journal of the American Statistical Association 1983:628-637.
33. Teahan W, Wen Y, McNab R, Witten I. A compression-based algorithm for Chinese word segmentation. Computational Linguistics 2000;26(3):375-393.
34. Tang J, Meng Z, Nguyen X, Mei Q, Zhang M. Understanding the limiting factors of topic modeling via posterior contraction analysis. 2014 Presented at: 31st International Conference on Machine Learning (ICML); June 21-26, 2014; Beijing, China.

Edited by G Eysenbach; submitted 14.03.15; peer-reviewed by S Santana; comments to author 27.07.15; revised version received 14.09.15; accepted 20.02.16; published 10.05.16.

Please cite as:
Hao H, Zhang K
The Voice of Chinese Health Consumers: A Text Mining Approach to Web-Based Physician Reviews
J Med Internet Res 2016;18(5):e108
URL: http://www.jmir.org/2016/5/e108/
doi:10.2196/jmir.4430
PMID:27165558

Original Paper

Activity Theory as a Theoretical Framework for Health Self-Quantification: A Systematic Review of Empirical Studies

Manal Almalki[1,2*], MInfTech; Kathleen Gray[1*], PhD; Fernando Martin-Sanchez[1,3*], PhD

[1]Health and Biomedical Informatics Centre, Melbourne Medical School, The University of Melbourne, Melbourne, Australia

[2]Faculty of Computer Science and Information Systems, Information Systems Department, Jazan University, Jazan, Saudi Arabia

[3]Healthcare Policy and Research, Weill Cornell Medicine, Cornell University, New York, NY, United States

[*]all authors contributed equally

Corresponding Author:
Manal Almalki, MInfTech
Health and Biomedical Informatics Centre
Melbourne Medical School
The University of Melbourne
Room 106, Level 1
202 Berkeley Street
Melbourne, 3010
Australia
Phone: 61 3 8344 8936
Fax: 61 3 90358873
Email: manal_almalki@hotmail.com

Abstract

Background: Self-quantification (SQ) is a way of working in which, by using tracking tools, people aim to collect, manage, and reflect on personal health data to gain a better understanding of their own body, health behavior, and interaction with the world around them. However, health SQ lacks a formal framework for describing the self-quantifiers' activities and their contextual components or constructs to pursue these health related goals. Establishing such framework is important because it is the first step to operationalize health SQ fully. This may in turn help to achieve the aims of health professionals and researchers who seek to make or study changes in the self-quantifiers' health systematically.

Objective: The aim of this study was to review studies on health SQ in order to answer the following questions: What are the general features of the work and the particular activities that self-quantifiers perform to achieve their health objectives? What constructs of health SQ have been identified in the scientific literature? How have these studies described such constructs? How would it be possible to model these constructs theoretically to characterize the work of health SQ?

Methods: A systematic review of peer-reviewed literature was conducted. A total of 26 empirical studies were included. The content of these studies was thematically analyzed using Activity Theory as an organizing framework.

Results: The literature provided varying descriptions of health SQ as data-driven and objective-oriented work mediated by SQ tools. From the literature, we identified two types of SQ work: work on data (ie, data management activities) and work with data (ie, health management activities). Using Activity Theory, these activities could be characterized into 6 constructs: users, tracking tools, health objectives, division of work, community or group setting, and SQ plan and rules. We could not find a reference to any single study that accounted for all these activities and constructs of health SQ activity.

Conclusions: A Health Self-Quantification Activity Framework is presented, which shows SQ tool use in context, in relation to the goals, plans, and competence of the user. This makes it easier to analyze issues affecting SQ activity, and thereby makes it more feasible to address them. This review makes two significant contributions to research in this field: it explores health SQ work and its constructs thoroughly and it adapts Activity Theory to describe health SQ activity systematically.

(J Med Internet Res 2016;18(5):e131) doi:10.2196/jmir.5000

KEYWORDS

activities of daily living; diagnostic self-evaluation; patient activation; patient participation; self-care; self-experimentation; self-management; user-computer interface; activity theory; human-computer interaction; self-quantification; self-tracking; personal informatics; quantified self; self-monitoring

Introduction

People can now measure for themselves their heart rate, sleep quality and quantity, mood, workouts, blood pressure, food consumed, quality of surrounding air—anything from mental, emotional, and physical to social and environmental aspects of daily life—because of advances in wearable sensors and apps, or for short, self-quantification (SQ) tools [1]. Examples of SQ tools are Fitbit for counting steps and tracking sleep, Adidas miCoach for tracking physical activities such as swimming and running, and Lumo Back for monitoring posture. Health SQ has the potential to induce changes in behaviors: according to one US survey [2], 69% of adults tracked themselves; 21% of the study population was using dedicated devices, and 46% stated that they had changed their behavior based on the collected data.

To achieve positive behavior changes, self-quantifiers need to undertake various activities in order to acquire, quantify, and aggregate data about aspects of their personal health, and translate these activities into activities of daily living such as eating healthily, maintaining a healthy weight, and being physically active [3-7]. These activities can be conceptualized as a form of doing work [8]. During this work, users interact with their personal SQ tools in order to collect and reflect on their data [4,5,7,8]. Thus, these tools mediate the SQ work [9-13].

The mediation principle here suggests "a structure for human-computer interaction that (...) the components of the structure should be not only the user and the computer but also the object the user is operating on through the computer application and the other people with whom the user is communicating" [14]. To come to an adequate understanding of this structure of human-computer interaction, deconstruction of the work or overall activity of tool use is needed [15]. This deconstruction can be achieved by breaking down the work into its contextual components or constructs [14-17].

The primary aim of this study was to review studies on health SQ in order to explore the self-quantifiers' work and activities and to answer the following questions: what constructs of health SQ work have been identified and examined in the relevant literature? How have these studies described such constructs? This review tests our hypothesis that health SQ can be characterized in two main complementary ways: as work on data and work with data.

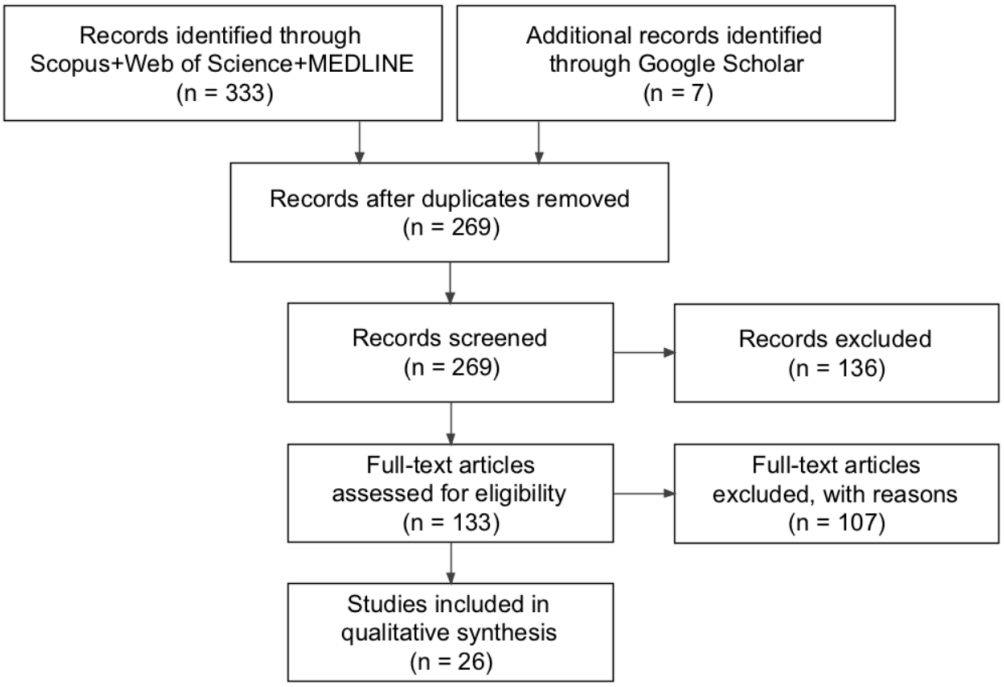
The secondary aim of this review was to model health SQ constructs, in order to better characterize the SQ work in health. Activity Theory (AT) appears to provide an appropriate conceptual basis to model health SQ work, because the unit of analysis in AT is the activity or work [16,18], within which 6 key constructs are recognized: subject, instruments, objectives, division of work, community, and plans and rules [14-16]. Therefore, this paper tests the adequacy of AT as a way to conceptualize health SQ.

Methods

Search Strategy

We searched Scopus, Web of Science, MEDLINE, and Google Scholar. We used the search strings "quantified self," "quantif*+self*," and "quantif*AND self." The search included studies published from January 2007 to October 2015. We included journals, conference proceedings, and papers written in English. The review was guided by the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement for reporting systematic reviews [19] (Figure 1).

Figure 1. Flow diagram for health self-quantification systematic review.



Inclusion and Exclusion Criteria

Twenty-six empirical studies met our inclusion criteria (see Multimedia Appendix 1). This sample was achieved through screening titles, abstracts, and keywords of 340 studies (Scopus 211, Web of Science 102, MEDLINE 20, and Google Scholar 7). A total of 71 studies were duplicated; of the remainder, we excluded 243 studies, in two rounds. In the first round, 136 studies were excluded because these studies either did not consider the use of SQ technologies for health self-management or did not investigate users in real-world contexts or daily life settings (rather than in controlled environment laboratories) where people individually collected, managed, and reflected on data. As a result of this round, 133 articles were eligible for the next round of full-text screening. In this round, exclusion criteria were as follows: studies provided conceptual knowledge but not empirical evidence, eliminating 78 articles; and studies focused on proposing a new solution and reporting only its technical specifications, eliminating 29 articles (see Multimedia Appendix 2) [20-122].

Characteristics and Quality Assessment of Included Studies

To assess the quality of the included studies, we followed the method recommended by Thomas and Harden [123].

Included studies in this review were coded by authors' names, year of publication, study aims, study design, and study main outcomes or results (Multimedia Appendix 1). Other characteristics noted for each study were as follows: the sampling frame (ie, source of data, selection of participants, recruitment methods, and consent); data collection methods (eg, interviews, questionnaires, and so on); the strategies used to ensure the reliability and validity of data collection methods; data analysis methods (eg, inductive content analysis, statistical quantitative analysis, and so on); and the strategies used to ensure the reliability and validity of data analysis methods.

Included studies in this review were then assessed according to 12 criteria, covering three main quality issues. Five criteria are related to the quality of the reporting of a study's aims and objectives, sampling frame, and data collection and analysis methods (eg, was there an adequate description of the sample used and the methods for how the sample was selected and recruited?). A further 4 criteria are related to the sufficiency of the strategies employed to establish the reliability and validity of data collection and analysis methods, and hence the validity of the findings. The final 3 criteria are related to the appropriateness of the study design (ie, appropriateness of the included study design considering our research aims). The first 9 criteria are suggested by Thomas and Harden [123], whereas the last 3 criteria, related to the appropriateness of the study design, are left open and can be decided by the researcher. In this review, these criteria are assumed to be met by setting the inclusion and exclusion criteria as explained in the previous section, and including only empirical studies for the thematic analysis. We found that all the included studies have reported adequate descriptions of the research aims and objectives, sampling frame, and data collection and analysis methods as well as good or some attempt to establish the reliability and validity of the data collection and analysis methods; thus, they met most of these quality criteria (see Multimedia Appendix 1).

Extracting Data from Included Studies and Thematic Analysis

To extract and synthesize themes about health SQ work, first we inductively coded the included studies [124,125]. Then, to investigate the constructs of health SQ work in the light of AT, we deductively coded the included studies into the following

themes [124,125]. Subject is a person engaged in an activity and using instruments in the course of this work [16]. Object (in the sense of "objective") is held by the subject and motivates activity, giving it a specific direction [126]. Division of work relates to the extent to which an activity involves collaboration and the sharing of tasks with others [16] or with tools [127]. Community is a group of people with whom the subject shares similar objectives [14,16]. Plans and rules are the norms and specifications of an activity that is undertaken by subjects to fulfill their objectives [16]. Open coding of the content of each study was done by looking at the language it used to explain and describe SQ and associating terms and synonyms with these themes. Both stages of analysis were undertaken by author MA, and then the preliminary results were critiqued jointly by authors KG and FMS, in several rounds of review, until all authors reached agreement on the interpretation of the data.

Results

Self-Quantification as Work "On Data" and "With Data"

We found that the literature provided a variety of perspectives on health SQ as data-driven work that users undertake to fulfill their health objectives [4,127,128]. It described how users interact with the SQ tools to define what aspects are relevant to their health conditions (eg, weight, sleep, blood pressure, and so on) [4,5]; set goals and track data about these health aspects for a period of time [7]; analyze the collected data to extract insights on health status [129,130]; adjust behaviors based on the insights and knowledge obtained from the analyzed data [130]; and control the adapted behaviors by sustaining the changes until the desired health outcomes are achieved [127]. We found that it was possible to categorize the overall content of the literature on health SQ activity or work in two ways: work on data and work with data. Table 1 maps these to the included studies. We found that the literature addressed working on data much more that working with data.

The first type, work on data, refers to activities that users carry out to manage their health SQ data. We found that the description of health SQ activities was not consistent across the reviewed literature, even when studies were based on the same human-computer interaction concepts. One study [4] stated that there are five key activities: preparation (eg, determining what information to collect and what collection tool to use), data collection, data integration (ie, combining data from multiple tools into one place), reflection (ie, analyzing data and looking at or exploring visualized information), and action-taking by the tools (eg, sending alerts to the user). However, another study [7] asserted that there are three key activities: goal setting and collecting data via using the tools (eg, BodyMedia—an armband for weight management—allows users to set goals toward more physical activity, lose weight, or stay the same); interpretation or reflection on the data (eg, analyzing data and calculating the time necessary to reach goals by the software); and providing feedback and coaching by the software (eg, Larklife—a wristband that tracks steps taken—will glow in a certain color when the users have not been moving for a while). Furthermore, we found that data collection and analysis activities received high attention, whereas hardly any study examined data handling activity (eg, data organizing, storing, and so on).

The second type, work with data, refers to activities that users carry out in using their data as the basis for actively managing their health status [8,10]. Some studies used the term "reflection" to imply this type of activities; to reflect upon data, users are assumed to look at their analyzed data, attempt to understand them, and act upon this understanding to improve or maintain health status (as in studies [4,7]). Other studies described this kind of interaction from the perspective of users' cognition in relation to one or more factors: building belief [12,131]; intentions for collecting data [131]; perception of the usefulness of using such tools in tracking health [132]; perception of the usefulness of objective data in making informed decisions [133], building knowledge or awareness about health and function status [5,7,130,134], in deciding what action should be taken on the generated data [8], the ability to act [132], and maintaining behavior or keeping track to cope with growing health problems or conditions [135]; and perception of themselves as good or bad self-quantifiers [8]. However, the definition of these factors is vague and questions remain open about how these factors are interrelated. We found that there is a need for further conceptualization to define and describe this kind of interaction between users and their health SQ data.

Constructs of Overall Health Self-Quantification Activity

Applying the concepts of AT to our analysis of included studies, we found that one study [8] mentioned all 6 AT constructs, although it did not examine their structure or the relationships among them. The remaining studies mentioned only 3 constructs—user, tools, and objectives. Table 1 maps the AT constructs mentioned in the included studies, and the following paragraphs give detailed examples of how we found these occurring in the literature.

Table 1. Mapping of themes to included studies in a systematic review of health self-quantification.

Themes	Article reference number																											%
	4	5	6	7	8	9	10	12	13	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	
Inductive themes																												
Work on data																												
	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	100
Work with data																												
	X	X		X	X		X	X				X	X	X	X	X	X	X									50
Deductive themes from AT[a]																												
Subject or user																												
	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	100
Instruments or tools																												
	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	100
Objectives																												
	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	100
Division of work																												
	X				X																						8
Community																												
	X	X			X	X	X	X	X		X	X	X			X	X	X		X				X	X	X	65
Plan and rules																												
			X		X					X																	11.5

[a]AT: Activity Theory.

Subject

A subject is a user who "tracks many kinds of data about themselves" [5]. They could be life hackers, data analysts, computer scientists, computer literate, early adopters, health enthusiasts, or productivity gurus [5]. They could also be self-experimenters and use such tools to conduct their own health studies and experiments [5,9,10,136]. They could be people who track themselves diligently and have relatively high technical and mathematical skills [5,8,9,130]. On the other hand, users could be more elderly, less educated, less affluent people, or people with one or more chronic conditions; such people could have more limited skills, less experience of using technology, or less access to SQ tools [8,130,137].

Instruments

Instruments are the tools that are used for health SQ. The description of tools in the reviewed studies was loose; tools were variously described as systems, devices, sensors, applications, mobile phone apps, Web-based applications, or software. Such variations might be assumed to be based on the characteristics and features of various technologies; however, this was not always the case. The term "system" at times was used to refer to a Web-based application, such as the prototype described by Li [4], and sometimes was used to refer to the health SQ device and its accompanying software application (such as in studies [138-140]). We note that the term "application" is not self-explanatory; there are many different types of applications such as mobile phone–based applications, Web-based applications, and applications associated with physical sensors (eg, the Fitbit app is associated with the Fitbit clip sensor and cannot be used fully without the sensor). Our analysis of this inconsistency shows a gap in the description of SQ tools and indicates the need for an agreed taxonomy.

Objectives

Objectives, in practical terms, are what users aim to achieve (eg, to increase exercise levels) [7]. Setting goals in relation to these objectives is enabled by the tools used [7]. The SQ tools convert these goals into a qualitative form of data (eg, activeness state, such as active or inactive) or a quantitative form of data (eg, number of steps per day) that can be used for illustrating the users' progress toward accomplishing what they want [4]. These data were described as indicators, health aspects, variables, metrics, parameters, and health factors, based on the researchers' view of the purpose of SQ work. For instance, the expressions "health indicators" and "health aspects" were used often to describe data obtained through tracking and personal analytics (such as in studies [4-6,12,13,128,135,141]), whereas health "variables," "metrics," "parameters," and "factors" were more associated with data produced as a result of self-experimentation (such as in studies [5,127,136,142]). In addition, these objectives could be described from a health perspective. One study classified these data into five categories

based on their health and well-being characteristics: body state (eg, physical and physiological), psychological state and traits, activities (eg, exercise, eating, sleeping), social interactions, and environmental and property states [128]; however, this classification is not adequate to account for all health-related aspects in each category. Our review indicates a need for a classification scheme for describing the SQ data in terms of the health objectives, in a comprehensive and systematic manner.

Division of Work

The literature showed that the division of work could be interpreted in two ways. In the first case, the SQ work can be technically divided between the users and the tools used [4]. For example, in data collection, the tools may offer manual or automatic data collection. In manual data collection the user needs to perform the required action (eg, logging food intake manually by using CalorieTracker app), whereas in the automatic mode the tools perform the required action (eg, counting steps automatically by Fitbit) [4].

In the second case, the work is divided between users and others (eg, friends, peers, and so on) [8]. Here, the users not only need to collect data and know numbers, but also need to assess their perception of their own status and achievements to gain a holistic understanding about themselves. To do so, they may share their data or experience with peers [6,13,143]. For example, sharing the exact steps taken in a week with a group of peers allows the person to compare their numbers with others and then accordingly evaluate their own personal activity level [9]. In another example, the users may share their findings from health self-experimentation with others who have run similar experiments in order to compare results and then confirm or disconfirm their own hypotheses [5,136]. In addition, some users go online and share their health concerns; for example, to discuss relevant aspects of their health conditions [143], about which we give further details in the next section.

Community

Community refers to the persons with whom the users opt to share their SQ experience. Many SQ tools support sharing data via social sharing features (eg, Fitbit allows users to share their data with groups or other individuals) [134]. Self-quantifiers can voluntarily share their data, results, and so on with others who could be friends, family members, relatives, partners, health care professionals [5,7-9,134,135,141], or peers on online and traditional health support groups (eg, PatientsLikeMe, CureTogether, Quantified Self groups, and so on) [5,137,143] or social networks (eg, Twitter, Facebook, and so on) [134,135,143]. By doing so they hope to get motivated, learn together, aggregate insights, compare results, compete or game, engage in teamwork, and so on. Also, in pursuit of supporting research, self-quantifiers may share their SQ experiences with medical researchers who conduct health-related studies at the population level (eg, crowd-sourcing studies over the Internet like Genomera, DIYgenomics, and so on) [10,137,143] or SQ technology researchers who may ask participants to test a certain technical aspect of a tool in natural settings [4,10,12].

Plan and Rules

Plan and rules in AT can refer to the method that users decide on to reach their health objectives. In the examined literature, these plans were described as styles of tracking [6] or data collection plans [127]; approaches of using personal data for health self-management [8]; or improvement plans that users design for improving their lifestyle and adjusting behavior based on the insight obtained from the previous tracking experience [127].

One study [6] suggested five methods: directive toward a goal, such as to either lose or maintain weight; documentary, to keep an eye on things but not to change them; diagnostic, to look for associations between health aspects (eg, tracking medication intake and diet to find out the cause of stomach problems); collecting rewards, to score points or register achievements; and fetishized, that is, driven by the interest in trying out new gadgets and technology. A further four methods were added by another study [8]: to take action (eg, tracking blood glucose to adjust diet or medication); to check progress toward goals but not necessarily for taking actions (eg, tracking cholesterol, and blood count values in anemia); to make sense of the health condition status (eg, checking glucose when users experience symptoms that they suspect indicate hypoglycemia, such as feeling light-headed); and to show logs to health care providers.

On the other hand, the rules of health SQ could be related simply to the regularity and frequency of tracking [4,8]. One study described how some users examine their data periodically for a holistic check on their own health, while others use their data frequently for real-time decisions about their behavior [8]. Another study described how users might collect data several times a day (eg, food consumption), once a day (eg, amount of sleep), several times a week (eg, exercise), or a few times a month (eg, symptoms) depending on their health needs and observations [4].

Discussion

In health SQ, people must undertake many different activities to transform their objectives into the desired outcome. Preparing, acquiring, organizing, maintaining, retrieving, and reviewing data in order to understand health status are related to data management (DM) [144]; therefore, health SQ can be described as a set of DM activities. However, people must undertake activities that go beyond DM in order to actually take initiative and responsibility for managing their health. This kind of work is related to health management (HM) [145-147]; therefore, health SQ must also be described as a set of HM activities. We could not find any study in the literature that identified all the AT constructs of these activities and examined the relations between them.

This limited current view makes it difficult to holistically investigate the effects of health SQ or systematically determine which “constructs” or components could be key to supporting or undermining users’ abilities to pursue their health objectives, and hence to achieve their desired health outcomes. Thus, a rigorous theoretical framework is needed, one that facilitates deeper, multi-aspect, and more systemic understanding of the

SQ activity or interaction between users and tools, highlights all the constructs of such activity, and maps them in a structured way [17,131]. This makes it easier to analyze issues affecting the SQ work or activity, and hence makes it more feasible to address them [17]. The following sections present a way to map the constructs identified throughout the literature, extending AT to develop a *Health Self-Quantification Activity Framework*.

Mapping All the Constructs of Health SQ Activity: Development of the Health Self-Quantification Activity Framework

In the AT system, the subject has a goal, tools, colleagues or friends, and rules when he or she is working on the goal to transform it into the intended outcome [15,16,148]. This goal is transformed into outcomes through a process of doing or activities [16]. This process is called the transformation process [16,148]. In this paper, this transformation process represents the work on data and with data that health self-quantifiers perform to manage their own data and health, respectively. In the light of AT, we can say that self-quantifiers have health objectives (eg, being an active or fit person). To attain such objectives, they use SQ tools and set SQ plans to track data in a certain way (eg, directive toward a goal), for example, to walk 10,000 steps per day. They interact with the SQ tools in order to collect, analyze, store, and share data with others as well as to gain knowledge on health, take actions on the generated data to protect or maintain health status, and so on. Thus, AT appears to provide an appropriate conceptual basis to describe health SQ work. Therefore, it is used as a foundation to develop the *Health Self-Quantification Activity Framework* (see Figure 2).

To illustrate how this framework fits with the AT conceptualization of work, the discussion next presents the reconciliation between the AT's 5 principles and the nature of health SQ practice. These principles constitute the general conceptual system of an activity, which are as follows: object-orientedness, tool mediation, the historical development of activity, the hierarchical structure of activity, and internalization/externalization [149]. Each of these principles will be discussed in turn.

Activity Theory asserts that the subject's work is an objective-oriented activity [126,150]. Objectives are sense-forming motives [126] where subjects' consciousness of the world around them is formed by their acting upon it [15]. In addition, an objective can be a motive-stimulus that stimulates a subject to engage in the activity until the desired goals are reached [126]. In health SQ, self-quantifiers track themselves for various purposes or objectives such as to improve health (eg, track blood glucose to hit the target range); to improve other aspects of well-being (eg, track time spent on doing things to be mindful); and to find new life experiences (eg, track heart rate for as long as possible to see what can be learned from it) [5]. Setting goals by using SQ tools in relation to these objectives helps users to collect the data necessary to attain their targets [5,8,13,137]. For example, in case of using Jawbone UP—a wearable wristband for tracking physical activities such as walking—when the users have not been moving a lot, Jawbone UP will start vibrating after being still for a certain time [7]. This helps the users to make sense of their health status (eg, I have not been active today), make a real-time decision (eg, do more walking now), or perform medium-term self-assessment (eg, if a person with diabetes is not feeling well or is inactive today, they may need to check their blood sugar) [8]. It also helps to illustrate one's progress toward reaching what they want; hence, it stimulates or motivates users to keep on working until their goals are achieved [4-7,13].

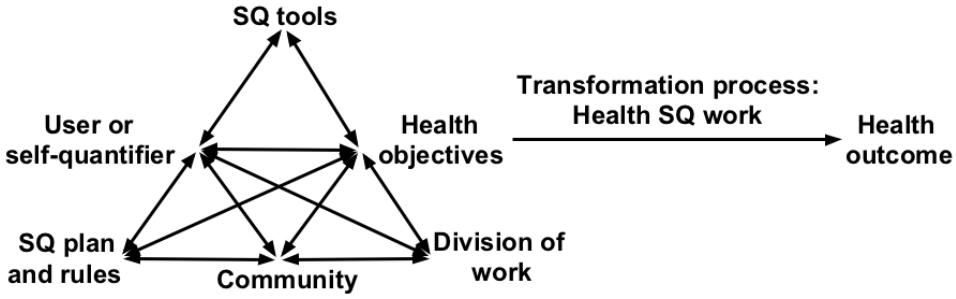
Activity Theory asserts that the subjects' work is mediated by tools [14-16]. Thus, tools could be at the same time both enabling and limiting. They may empower the users during their work to attain their goals, or they may restrict such interaction [15]. Either effect may be observed in health SQ. Self-quantification activities are mediated by the users' tools [9-13]. As an example of mixed effects, a study about using SQ tools to measure progress by athletes [9] found that if the generated data or scores are below what was expected (ie, because of limitation in data accuracy that users are not aware of), their confidence and how they perceive themselves as athletic individuals may be distorted; consequently, their performance could decrease and eventually they might stop analyzing or even collecting such data.

Activity Theory states that the objective-oriented activity has a history of its own from which the human mind develops an understanding [150]. The historical analysis of the activity is often needed to understand the recent situation [15]. In health SQ, the advances in computational analysis of SQ tools make building a history of work possible, which is a major facilitator to understand current health status and obtain self-knowledge [129,130,134,142]. Data generated from using SQ tools are also beneficial for people to evaluate their future health status. Providing a history of the collected data and detecting trends over time can help users to not only quantify the current health and function status, but also calibrate expectations for upcoming activities based on previous experience. Through evaluating current performance against past performance, they may compare their health status at different times in the year [9,141].

Activity Theory differentiates between activity, actions, and operations to offer a hierarchical structure of activity. Achieving the subjects' objectives requires them to go through a series of activities [15,16,18]. Each of these activities is composed of actions, and actions are composed of operations [16]. In health SQ, users undertake various DM and HM activities in pursuit of achieving their health objectives. These activities are composed of actions and operations [1]. We can take as an example the data collection activity, which is one of the DM activities: in the case of quantifying the walking habit, it could be composed of several actions such as setting goals, wearing the tools, going for a walk, and so on that are required for generating data or measuring performance [1]. However, the conceptualization of activities and actions in the examined literature is vague; thus, the boundary between them is not clear. This could be one of the reasons for inconsistencies in describing the DM activities, as discussed previously. For example, goal setting was considered a key activity by De Maeyer and Jacobs [7], whereas data collection is an action that constitutes this activity. In contrast, Li [4] considered data collection as a key activity, whereas goal setting is an action that may be a part of this activity.

Activity Theory asserts that any activity has a recursive structure in which feedback that evaluates performance is decisive [150]. On the basis of this feedback, the subject's internal mental or cognitive actions are formed and this is called internalization. The opposite of the internalization process is the externalization process where internal cognitive actions are transformed into external actions. These two processes demonstrate that cognitive activity is tightly interconnected with external objective-practical activity [150]. In health SQ, data generated from SQ tools may provide decisive feedback about the person's current health status because they are neutral and not intuitive or emotional [3,8,11,133]. Such data offer an opportunity for users to obtain self-knowledge about health status, and based on this knowledge they decide on what actions they need to take [4,7,9,13]. However, our analysis of the literature shows that these internalization/externalization processes have not been fully investigated yet, and the transition between these two processes remains open for examination. Therefore, a more detailed discussion of these processes cannot be provided in this paper and will be taken up in future work.

Figure 2. Health Self-Quantification Activity Framework. SQ: self-quantification.



Using the Health Self-Quantification Activity Framework to Investigate Health SQ Work

Many self-quantifiers have been frustrated in accomplishing their health objective [4-10]. They gradually lose their initial motivation and may stop tracking, and thus may not achieve the desired health outcomes [1,4-10,128]. To examine what could cause such a situation in a more systematic and structured way, the *Health Self-Quantification Activity Framework* puts SQ tool use into context, in relation to the goals, plans, and competence of the user. In the following paragraphs, examples of how a person could fail to achieve her or his goals are provided in the light of the health SQ constructs as follows.

"Users" need to have high information and communication technology (ICT) skills [6] and mathematical skills [9] in order to successfully use the SQ tools to track health. However, not everyone is so skilled. A study shows that few people in Europe have high ICT skills (only 29% of the European population from 28 countries) [151]. Therefore, factoring in the users' technical and mathematical competence levels (eg, low, moderate, or high) could lead to better understanding their health outcomes from health SQ [152,153].

"SQ tools" have a major influence over users. A study found that "users were emotionally involved with their tools, and saw this more as their external conscious to see the details on their physical activity, food and sleep patterns and act on it where necessary" [7]. However, any issues that arise when these tools are in use may increase the time and cognitive effort needed to resolve these issues and to gain self-awareness [4,5,7,154]. For example, if many types of health indicators are of interest, finding a single tool that can support acquisition of most of these data is difficult [4-7,12,138], thus, the users may decide not to collect all the data that would be useful to successfully manage health [4]; if they proceed, they may find themselves using multiple tools, which increases the person's time and efforts to integrate, analyze, and learn from the collected data [4,5]; and if the same data are tracked by different tools, when combined, data may be inconsistent, giving rise to concerns about data accuracy [139]. Therefore, users may lose their trust in these measurements or may stop tracking [4,10,139,141].

"The division of work"—for example, dividing data collection work between the user and tools in a form of manual or automatic data collection [4,6,10]—may be cumbersome, and this can impede the user's motivation to collect data of interest or pursue the SQ practice [7]. For example, one study [4] found that self-quantifiers were eager to log their food intake manually at the beginning, but this decreased because as time passed they found it very time consuming.

"Community" is a very real experience that is enabled by health SQ: some users go online and share their data or results about their health status with others who could be friends, family, or strangers, in order to observe, compare, compete, and so on [143]. However, sharing this kind of information has its own set of problems. For example, it could expose the users to privacy breaches and different kinds of discrimination, including discrimination in insurance and other financial dealings [143]. In addition, if the users' results appeared lower when compared with others, they might consequently develop a low opinion of themselves and thus they might stop tracking their performance [9,137].

"SQ plans and rules" are different for each person, and most SQ systems do not consider this variation. For example, using SQ tools for conducting self-experimentation raises many issues [5,10,136]. When the individuals have personal hypotheses (eg, if I drink sugar water over a period of 8 weeks, then I will lose weight), they want to design their tracking plan to either prove or disprove their hypotheses. However, the lack of integration of single-case experimental design standards into health SQ

systems leads to insufficient scientific rigor in data collection and analysis [5,136]. Thus, with current SQ tools, testing one's thinking about personal weight loss would be challenging and potentially misleading.

Conclusions

To summarize, health SQ is data-driven and objective-oriented work that is mediated by tools. This work is composed of two types of activities: DM activities (ie, work on data) and HM activities (ie, work with data). These activities comprise 6 constructs (ie, users, SQ tools, health objectives and goals, division of work, community, and SQ plans and rules). Understanding health SQ work, activities, and constructs is important because it is the first step to operationalize health SQ fully. This may in turn help to achieve the aims of health professionals and researchers who seek to make or study changes in the self-quantifiers' health systematically. This review makes two significant contributions to research in this field: it explores health SQ work and its constructs thoroughly and it adapts AT to describe health SQ activity systematically.

However, many gaps exist in the literature (eg, inconsistency in describing tools, lack of a comprehensive view of health objectives, and vagueness in describing DM and HM activities) that need further investigation. To fill these gaps, our ongoing work is leading toward developing a taxonomy that accounts for various aspects of SQ tools' functionality and characterizations [1]; a conceptual model and classification schema for explaining data or health aspects [155]; and a model that describes the key DM and HM activities that are necessary for health self-management [1,154]. These are critical contributions to establishing a holistic and rigorous theoretical framework within which to understand and improve health SQ activity.

Acknowledgments

Funding for MA's PhD research is supported by a scholarship from Jazan University and a top-up scholarship from the Melbourne Networked Society Institute at the University of Melbourne.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Included studies, their characteristics, and quality assessment.

[PDF File (Adobe PDF File), 84KB - jmir_v18i5e131_app1.pdf]

Multimedia Appendix 2

Excluded studies and reasons for exclusion.

[PDF File (Adobe PDF File), 34KB - jmir_v18i5e131_app2.pdf]

References

1. Almalki M, Gray K, Sanchez F. The use of self-quantification systems for personal health information: big data management activities and prospects. Health Information Science and Systems 2015;3(Suppl 1):S1. [doi: 10.1186/2047-2501-3-S1-S1]
2. Fox S, Duggan M. Tracking for health. Washington: Pew Research Center; 2013 Jan 28. URL: http://www.pewinternet.org/Reports/2013/Tracking-for-Health.aspx [accessed 2016-05-14] [WebCite Cache ID 6Zv4xkIxB]
3. Lupton D. SSRN: Social Science Research Network. 2014 Dec 04. You are your data: self-tracking practices and concepts of data URL: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2534211 [accessed 2016-05-11] [WebCite Cache ID 6hRKVBvGq]
4. Li I. Personal informatics and context: using context to reveal factors that affect behavior. Journal of Ambient Intelligence and Smart Environments 2012;4(1):71-72. [doi: 10.3233/AIS-2011-0130]
5. Choe E, Lee N, Lee B, Pratt W, Kientz J. Understanding quantified-selfers' practices in collecting and exploring personal data. In: CHI '14 Proceedings of the SIGCHI conference on Human factors in computing systems. New York: ACM; 2014 Presented at: SIGCHI conference on Human factors in computing systems; April 26 - May 1, 2014; Toronto, Ontario p. 1143-1152. [doi: 10.1145/2556288.2557372]
6. Rooksby J, Rost M, Morrison A, Chalmers M. Personal tracking as lived informatics. In: CHI '14: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; Toronto, Ontario, Canada. New York: ACM; 2014 Presented at: SIGCHI Conference on Human Factors in Computing Systems; April 26 - May 1, 2014; Toronto, Ontario p. 1163-1172. [doi: 10.1145/2556288.2557039]
7. De Maeyer C, Jacobs A. Sleeping with technology - designing for personal health. In: Matsumura N, editor. Shikakeology - designing triggers for behavior change: Papers from the 2013 AAAI Spring Symposium. Palo Alto, CA: AAAI Press; 2013:11-16.

8. Ancker JS, Witteman HO, Hafeez B, Provencher T, Van de Graaf Mary, Wei E. “You Get Reminded You're a Sick Person”: Personal Data Tracking and Patients With Multiple Chronic Conditions. J Med Internet Res 2015 Aug;17(8):e202 [FREE Full text] [doi: 10.2196/jmir.4209] [Medline: 26290186]
9. Lee VR, Drake J. Digital Physical Activity Data Collection and Use by Endurance Runners and Distance Cyclists. Tech Know Learn 2013 Apr 27;18(1-2):39-63. [doi: 10.1007/s10758-013-9203-3]
10. FIORE-SILFVAST B, NEFF G. What we talk about when we talk data: Valences and the social performance of multiple metrics in digital health. Ethnographic Praxis in Industry Conference Proceedings 2014 Feb 11;2013(1):74-87. [doi: 10.1111/j.1559-8918.2013.00007.x]
11. Saunders W, Krynicki F, Sugarman V. Sisyphorest: maintenance goal support by responding to trends. In: CHI EA '14 CHI '14 Extended Abstracts on Human Factors in Computing Systems. New York: ACM; 2014 Presented at: CHI '14 ACM CHI Conference on Human Factors in Computing Systems; April 26 - May 1, 2014; Toronto, Ontario p. 215-220. [doi: 10.1145/2559206.2580934]
12. Kim J. A qualitative analysis of user experiences with a self-tracker for activity, sleep, and diet. Interact J Med Res 2014 Mar;3(1):e8 [FREE Full text] [doi: 10.2196/ijmr.2878] [Medline: 24594898]
13. Gimpe H, Nißen M, Görlitz R. Quantifying the quantified self: A study on the motivations of patients to track their own health. In: International Conference on Information Systems ICIS 2013. Milan, Italy: Association for Information Systems (AIS); 2013 Presented at: 34th International Conference on Information Systems ICIS; 2013; Milan.
14. Kaptelinin V. Activity theory: implications for human-computer interaction. In: Nardi B, editor. Context and consciousness: activity theory and human-computer interaction. Cambridge, MA: MIT Press; 1995:103-116.
15. Kuutti K. Activity theory as a potential framework for human-computer interaction research. In: Nardi B, editor. Context and consciousness: activity theory and human-computer interaction. Cambridge, MA: MIT Press; 1996:17-44.
16. Wilson TD. Activity theory and information seeking. Ann Rev Info Sci Tech 2008;42(1):119-161. [doi: 10.1002/aris.2008.1440420111]
17. van Gemert-Pijnen Julia E W C, Nijland N, van LM, Ossebaard HC, Kelders SM, Eysenbach G, et al. A holistic framework to improve the uptake and impact of eHealth technologies. J Med Internet Res 2011 Dec;13(4):e111 [FREE Full text] [doi: 10.2196/jmir.1672] [Medline: 22155738]
18. Nardi B. Studying context: A comparison of activity theory, situated action models, and distributed cognition. In: Nardi B, editor. Context and consciousness: Activity theory and human-computer interaction. Cambridge, MA: MIT Press; 1996:69-102.
19. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. BMJ 2009;339:b2535 [FREE Full text] [Medline: 19622551]
20. Andrieu B. L’osmose technique avec son corps viv@nt : une auto-santé connectée du patient immersif (Technical osmosis with the living body: Health self-tracking of the immersive patient). L'Évolution Psychiatrique 2016 Apr;81(2):253-265. [doi: 10.1016/j.evopsy.2015.07.004]
21. Banos O, Amin M, Khan W, Afzel M, Ahmad M, Ali M. An innovative platform for person-centric health and wellness support. In: Ortuno F, Rojas I, editors. Bioinformatics and biomedical engineering. Cham, Switzerland: Springer International Publishing; 2015:131-140.
22. Barricelli B, Valtolina S. Designing for end-user development in the internet of things. In: Diaz P, Pipek V, Ardito C, Jensen C, Aedo I, Boden A, editors. End-User Development. Cham, Switzerland: Springer International Publishing; 2015:9-24.
23. Chamberlain A, Schraefel M, Poole E, Munson S, Danis C, Churchill E. Moving beyond e-health and the quantified self: the role of CSCW in collaboration, community and practice for technologically-supported proactive health and wellbeing. In: CSCW'15 Companion: Proceedings of the 18th ACM Conference Companion on Computer Supported Cooperative Work & Social Computing. New York: the role of CSCW in collaboration, community and practice for technologically-supported proactive health and wellbeing. Proceedings of the ACM Conference on Computer Supported Cooperative Work, CSCW; 2015 Presented at: CSCW '15: 18th ACM Conference Companion on Computer Supported Cooperative Work & Social Computing; 14-18 March, 2015; Vancouver, BC, Canada p. 273-276. [doi: 10.1145/2685553.2685555]
24. Davies N, Friday A, Clinch S, Sas C, Langheinrich M, Ward G, et al. Security and Privacy Implications of Pervasive Memory Augmentation. IEEE Pervasive Comput 2015 Jan;14(1):44-53. [doi: 10.1109/MPRV.2015.13]
25. Doryab A, Frost M, Faurholt-Jepsen M, Kessing LV, Bardram JE. Impact factor analysis: combining prediction with parameter ranking to reveal the impact of behavior on health outcome. Pers Ubiquit Comput 2014 Sep 21;19(2):355-365. [doi: 10.1007/s00779-014-0826-8]
26. Dudley JT, Listgarten J, Stegle O, Brenner SE, Parts L. Personalized medicine: from genotypes, molecular phenotypes and the quantified self, towards improved medicine. Pac Symp Biocomput 2015:342-346 [FREE Full text] [Medline: 25592594]
27. Forsdyke DR. Doctor-scientist-patients who barketh not: the quantified self-movement and crowd-sourcing research. J Eval Clin Pract 2015 Dec;21(6):1024-1027. [doi: 10.1111/jep.12425] [Medline: 26201555]
28. Giones-Valls A, Giones F. Quantifying the self to live through data: big data applied to the personal sphere. BiD: textos universitaris de biblioteconomia i documentació 2015(34):3 (forthcoming) [FREE Full text] [doi: 10.1344/BiD2015.34.16]

29. Hachem S, Mathioudakis G, Pathak A, Issarny V, Bhatia R. Sense2Health a quantified self application for monitoring personal exposure to environmental pollution. In: Proceedings of 4th International Conference on Sensor Networks, SENSORNETS.: Scitepress; 2015 Presented at: 4th International Conference on Sensor Networks, SENSORNETS 2015; February 11-13, 2015; Angers, Loire Valley, France p. 36-44. [doi: 10.5220/0005332100360044]
30. Hood L, Lovejoy JC, Price ND. Integrating big data and actionable health coaching to optimize wellness. BMC Med 2015;13:4 [FREE Full text] [doi: 10.1186/s12916-014-0238-7] [Medline: 25575752]
31. Issa H, Shafaee A, Agne S, Baumann S, Dengel A. Evaluation of Fitbit One, Jawbone Up and Nike+ Fuelband based on Amazon.com customer reviews. In: Proceedings of the 1st International Conference on Information and Communication Technologies for Ageing Well and e-Health.: SciTePress; 2015 Presented at: ICT4AgeingWell 2015; May 20-22, 2015; Lisbon p. 171-179. [doi: 10.5220/0005447401710179]
32. Janssen CP, Gould SJ, Li SY, Brumby DP, Cox AL. Integrating knowledge of multitasking and interruptions across different perspectives and research methods. International Journal of Human-Computer Studies 2015 Jul;79:1-5. [doi: 10.1016/j.ijhcs.2015.03.002]
33. Jones DH. self-archiving from Christian Boltanski to lifelogging. Archives and Records 2015 Feb 26;36(1):29-41. [doi: 10.1080/23257962.2015.1010149]
34. Lavalliere M, Arezes P, Burstein A, Coughlin J. The quantified-self and wearable technologies in the workplace: implications and challenges for their implementations. In: Sho2015: International Symposium on Occupational Safety and Hygiene. 2015 Presented at: Sho: International Symposium on Occupational Safety and Hygiene; February 12-13, 2015; Guimarães, Portugal p. 161-163 URL: http://hdl.handle.net/1822/38579 [WebCite Cache ID 6hd16IPgg]
35. Lupton D, Jutel A. 'It's like having a physician in your pocket!' A critical analysis of self-diagnosis smartphone apps. Soc Sci Med 2015 May;133:128-135. [doi: 10.1016/j.socscimed.2015.04.004] [Medline: 25864149]
36. Lupton D. Quantified sex: a critical analysis of sexual and reproductive self-tracking using apps. Cult Health Sex 2015 Jun;17(4):440-453. [doi: 10.1080/13691058.2014.920528] [Medline: 24917459]
37. Majmudar MD, Colucci LA, Landman AB. The quantified patient of the future: Opportunities and challenges. Healthc (Amst) 2015 Sep;3(3):153-156. [doi: 10.1016/j.hjdsi.2015.02.001] [Medline: 26384227]
38. Ohlin F, Olsson C. Intelligent computing in personal informatics: Key design considerations. In: Proceedings of the 20th International Conference on Intelligent User Interfaces. New York: ACM; 2015 Presented at: IUI '15 - 20th International Conference on Intelligent User Interfaces; 29 March - 1 April, 2015; Atlanta, GA p. 263-274. [doi: 10.1145/2678025.2701378]
39. Olivier B, Vilcocq-Merjagnan C. Online administration of a quantified self-questionnaire for elderly people: a user satisfaction survey. J Am Geriatr Soc 2015 Jan;63(1):194-195. [doi: 10.1111/jgs.13222] [Medline: 25597579]
40. Picard R, Wolf G. Guest Editorial Sensor Informatics and Quantified Self. IEEE J. Biomed. Health Inform 2015 Sep;19(5):1531. [doi: 10.1109/JBHI.2015.2462372]
41. Rüping S. [Big data in medicine and healthcare]. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 2015 Aug;58(8):794-798. [doi: 10.1007/s00103-015-2181-y] [Medline: 26063521]
42. Thornquist E, Kirkengen AL. The quantified self: closing the gap between general knowledge and particular case? J Eval Clin Pract 2015 Jun;21(3):398-403. [doi: 10.1111/jep.12239] [Medline: 25266335]
43. Tory M, Carpendale S. Personal Visualization and Personal Visual Analytics [Guest editors' introduction]. IEEE Comput. Grap. Appl 2015 Jul;35(4):26-27. [doi: 10.1109/MCG.2015.88]
44. Van den Bulck Jan. Sleep apps and the quantified self: blessing or curse? J Sleep Res 2015 Apr;24(2):121-123 [FREE Full text] [doi: 10.1111/jsr.12270] [Medline: 25558955]
45. Williams SJ, Coveney C, Meadows R. 'M-apping' sleep? Trends and transformations in the digital age. Sociol Health Illn 2015 Sep;37(7):1039-1054. [doi: 10.1111/1467-9566.12283] [Medline: 26182959]
46. Zandi AS, Boudreau P, Boivin DB, Dumont GA. Identification of scalp EEG circadian variation using a novel correlation sum measure. J Neural Eng 2015 Oct;12(5):056004. [doi: 10.1088/1741-2560/12/5/056004] [Medline: 26246488]
47. Appelboom G, LoPresti M, Reginster J, Sander CE, Dumont Emmanuel P L. The quantified patient: a patient participatory culture. Curr Med Res Opin 2014 Dec;30(12):2585-2587. [doi: 10.1185/03007995.2014.954032] [Medline: 25118077]
48. Becker BW. The Quantified Self: Balancing Privacy and Personal Metrics. Behavioral & Social Sciences Librarian 2014 Nov 10;33(4):212-215. [doi: 10.1080/01639269.2014.964595]
49. Brennan C, McCullagh P, Lightbody G, Galway L, Trainor D. Quantifying brain activity for task engagement. In: Proceedings of the IEEE International Conference on Bioinformatics and Biomedicine, BIBM.: IEEE; 2014 Presented at: IEEE International Conference on Bioinformatics and Biomedicine, BIBM; November 2 - 5, 2014; Belfast p. 8-11. [doi: 10.1109/BIBM.2014.6999282]
50. Bushhousen E. The Quantified Self movement and hospital librarians. Journal of Hospital Librarianship 2014 Feb 12;14(1):88-93. [doi: 10.1080/15323269.2014.861705]
51. Caldwell T. The quantified self: a threat to enterprise security? Computer Fraud & Security 2014 Nov;2014(11):16-20. [doi: 10.1016/S1361-3723(14)70549-4]
52. Calikli G, Andersen M, Bandara A, Price B, Nuseibeh B. Personal informatics for non-geeks: lessons learned from ordinary people. In: Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct

Publication. 2014 Presented at: UbiComp '14: 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing; September 13 - 17, 2014; Seattle, WA p. 683-686. [doi: 10.1145/2638728.2641316]
53. Cena F, Likavec S, Rapp A, Deplano M, Marcengo A. Ontologies for quantified self: a semantic approach. In: Hypertext 2014 extended proceedings of the 25th ACM Hypertext and Social Media conference.: CEUR-WS; 2014 Presented at: Hypertext 2014; September 1 - 4, 2014; Santiago, Chile URL: http://ceur-ws.org/Vol-1210/LQS_01.pdf [WebCite Cache ID 6hd1KWwGg]
54. Clifford GD, Gederi E. Out of touch : from audio recordings to phone apps to mattress sensors, noncontact systems offer a less cumbersome way to monitor sleep. IEEE Pulse 2014 Sep;5(5):19-21. [doi: 10.1109/MPUL.2014.2339253] [Medline: 25437469]
55. Cuttone A, Larsen J. The long tail issue in large scale deployment of personal informatics. In: Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication. 2014 Presented at: UbiComp '14 - ACM International Joint Conference on Pervasive and Ubiquitous Computing; September 13 - 17, 2014; Seattle, WA p. 691-694. [doi: 10.1145/2638728.2641319]
56. Cuttone A, Petersen M, Larsen J. Four data visualization heuristics to facilitate reflection in personal informatics. In: Stephanidis C, Antona M, editors. Universal access in human-computer interaction. Design for all and accessibility practice. Cham, Switzerland: Springer International Publishing; 2014:541-552.
57. De Croon R, De Buyser T, Klerkx J, Duval E. Applying a user-centered, rapid-prototyping methodology with quantified self: A case study with triathletes. In: Proceedings of IEEE International Conference on Bioinformatics and Biomedicine (BIBM). 2014 Presented at: IEEE International Conference on Bioinformatics and Biomedicine (BIBM); November, 2 - 5, 2014; Belfast p. 12-15. [doi: 10.1109/BIBM.2014.6999283]
58. Duarte F, Lourenço A, Abrantes A. Classification of physical activities using a smartphone: Evaluation study using multiple users. In: Proceedings of the 2013 Conference on Electronics, Telecommunications and Computers.: Elsevier; 2014 Presented at: CETC 2013 - Conference on Electronics, Telecommunications and Computers - CETC . Procedia Technology; December 5 - 6, 2013; Lisbon p. 239-247. [doi: 10.1016/j.protcy.2014.10.234]
59. Gastaldi M. Integration of mobile, big data, sensors, and social media: impact on daily life and business. In: Proceedings of the 2014 IST-Africa Conference and Exhibition.: IEEE; 2014 Presented at: 2014 IST-Africa Conference and Exhibition; May 7 - 9, 2014; Le Meridien Ile Maurice p. 1-10. [doi: 10.1109/ISTAFRICA.2014.6880670]
60. Gaunt K, Nacsa L, Penz M. Baby Lucent: pitfalls of applying quantified self to baby products. In: CHI '14 Extended Abstracts on Human Factors in Computing Systems. New York: ACM; 2014 Presented at: CHI '14: Human Factors in Computing Systems; April 26 - May 1, 2014; Toronto, Ont p. 263-268. [doi: 10.1145/2559206.2580937]
61. Gurrin C, Smeaton AF, Doherty AR. LifeLogging: Personal Big Data. FNT in Information Retrieval 2014;8(1):1-125. [doi: 10.1561/1500000033]
62. Harrison D, Marshall P, Berthouze N, Bird J. Tracking physical activity: problems related to running longitudinal studies with commercial devices. In: Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication. New York: ACM; 2014 Presented at: UbiComp '14 - 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing; September 13 - 17, 2014; Seattle, WA p. 699-702. [doi: 10.1145/2638728.2641320]
63. Hirsch JA, James P, Robinson Jamaica R M, Eastman KM, Conley KD, Evenson KR, et al. Using MapMyFitness to Place Physical Activity into Neighborhood Context. Front Public Health 2014 Mar;2:19 [FREE Full text] [doi: 10.3389/fpubh.2014.00019] [Medline: 24653982]
64. Hogenboom A, Smit I, Kröse B. Demonstration: a digital coach for self-tracking athletes. In: Haptics: Neuroscience, Devices, Modeling, and Applications - Proceedings of the 9th International Conference on EuroHaptics. Berlin: Springer Verlag; 2014 Presented at: EuroHaptics 2014; June 24 - 26, 2014; Versailles p. 430-431. [doi: 10.1007/978-3-662-44196-1]
65. Huang S, Sano A, Kwan C. The moment: a mobile tool for people with depression or bipolar disorder. In: Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication. 2014 Presented at: UbiComp '14 - ACM International Joint Conference on Pervasive and Ubiquitous Computing; September 13 - 17, 2014; Seattle, WA p. 235-238. [doi: 10.1145/2638728.2638784]
66. Jain R, Jalali L. Objective Self. IEEE MultiMedia 2014 Oct;21(4):100-110. [doi: 10.1109/MMUL.2014.63]
67. Jones C. Tell all. Innovations in Pharmaceutical Technology. 2014. (50) p. 14-17 URL: http://www.iptonline.com/synopsis.asp?cat=1&article=986 [WebCite Cache ID 6hd1TdO09]
68. Jordan M, Pfarr N. MDDI. From data to decisions-the path to a meaningful quantified self medical device and diagnostic industry news products and suppliers; MDDI 2014 URL: http://www.mddionline.com/article/data-decisions-path-meaningful-quantified-self [accessed 2015-07-12] [WebCite Cache ID 6ZxE7jXYh]
69. Keary A, Walsh P. How affective computing could complement and advance the quantified self. In: Proceedings of the Proceedings of the IEEE International Conference on Bioinformatics and Biomedicine, BIBM.: IEEE; 2014 Presented at: IEEE International Conference on Bioinformatics and Biomedicine, BIBM; November 2 - 5, 2014; Belfast p. 24-31. [doi: 10.1109/BIBM.2014.6999285]

70. Khorakhun C, Bhatti S. Wellbeing as a proxy for a mHealth study. In: Proceedings of the IEEE International Conference on Bioinformatics and Biomedicine, BIBM.: IEEE; 2014 Presented at: IEEE International Conference on Bioinformatics and Biomedicine, BIBM; May 2 - 5, 2014; Belfast p. 32-39. [doi: 10.1109/BIBM.2014.6999286]
71. Kido T, Swan M. Know Thyself: data driven self-awareness for understanding our unconsciousness behaviors. In: Big Data Becomes Personal: Knowledge into Meaning: Papers from the AAAI Spring Symposium.: AAAI; 2014 Presented at: AAAI Spring Symposium; March 24 - 26, 2014; Palo Alto, CA.
72. Lagus K. Looking at our data-perspectives from mindfulness apps and quantified self as a daily practice. In: Proceedings of the IEEE International Conference on Bioinformatics and Biomedicine, BIBM.: IEEE; 2014 Presented at: IEEE International Conference on Bioinformatics and Biomedicine, BIBM; November 2 - 5, 2014; Belfast p. 40-43. [doi: 10.1109/BIBM.2014.6999287]
73. Laundav D, Jensen C, Baekgaard P, Petersen M, Larsen J. Your heart might give away your emotions. In: Proceedings of the IEEE International Conference on Multimedia and Expo Workshops.: IEEE; 2014 Presented at: 2014 IEEE International Conference on Multimedia and Expo Workshops, ICMEW; July 14 - 18, 2014; Chengdu p. 1-6. [doi: 10.1109/ICMEW.2014.6890662]
74. Lee V, Briggs M. Lessons learned from an initial effort to bring a quantified self "meetup" experience to a new demographic. In: Proceedings of the ACM International Joint Conference on Pervasive and Ubiquitous Computing adjunct publication.: ACM; 2014 Presented at: UbiComp '14 - ACM International Joint Conference on Pervasive and Ubiquitous Computing; September 13 - 17, 2014; Seattle, WA p. 707-710. [doi: 10.1145/2638728.2641321]
75. Lingg E, Leone G, Spaulding K, B'Far R. Cardea: Cloud based employee health and wellness integrated wellness application with a wearable device and the HCM data store. In: Proceedings of the 2014 IEEE World Forum on Internet of Things.: IEEE; 2014 Presented at: WF-IoT 2014; March 6 - 8, 2014; Seoul p. 265-270. [doi: 10.1109/WF-IoT.2014.6803170]
76. Lupton D. Health promotion in the digital era: a critical commentary. Health Promot Int 2015 Mar;30(1):174-183. [doi: 10.1093/heapro/dau091] [Medline: 25320120]
77. Lupton D. Self-tracking modes: Reflexive self-monitoring and data practices. In: SSRN Journal.: SSRN; 2014 Presented at: Imminent Citizenships: Personhood and Identity Politics in the Informatic Age; August 27, 2014; Canberra p. 1-19 URL: http://ssrn.com/abstract=2483549 [WebCite Cache ID 6hd1eMj1s] [doi: 10.2139/ssrn.2483549]
78. Marcengo A, Buriano L, Geymonat M. Specch.io: a personal QS mirror for life patterns discovery and “self” reshaping. In: Stephanidis C, Antona M, editors. Universal access in human-computer interaction. Design for all and accessibility practice. Cham, Switzerland: Springer International Publishing; 2014:215-226.
79. Maturo A. Fatism, self-monitoring and the pursuit of healthiness in the time of technological solutionism. Italian Sociological Review 2014;4(2):157-171. [doi: 10.13136/isr.v4i2.80]
80. Maturo A. M-health and quantified self: developments, potential and risks. Salute e Societa 2014;3(13):161-170.
81. Meyer J, Simske S, Siek K, Gurrin C, Hermens H. Beyond quantified self: data for wellbeing. In: CHI '14 Extended Abstracts on Human Factors in Computing Systems. 2560469: ACM; 2014 Presented at: CHI '14 Human Factors in Computing Systems; April 26 - May 1, 2014; Toronto, Ont p. 95-98. [doi: 10.1145/2559206.2560469]
82. Nafus D, Sherman J. Big data, big questions| this one does not go up to 11: the quantified self movement as an alternative big data practice. International journal of communication. 2014. p. 1784-1794 URL: http://ijoc.org/index.php/ijoc/article/view/2170 [WebCite Cache ID 6hd1ic0VI]
83. Nikolic-Popovic J, Goubran R. Towards increased usability of noisy ECG signals in HRV-based classifiers. In: Proceedings of the IEEE International Symposium on Medical Measurements and Applications, MeMeA.: IEEE; 2014 Presented at: IEEE International Symposium on Medical Measurements and Applications, MeMeA; June 11 - 12, 2014; Lisbon p. 1-1. [doi: 10.1109/MeMeA.2014.6860125]
84. Prince JD. The Quantified Self: operationalizing the quotidien. Journal of Electronic Resources in Medical Libraries 2014 May 30;11(2):91-99. [doi: 10.1080/15424065.2014.909145]
85. Rapp A, Cena F. Self-monitoring and technology: challenges and open issues in personal informatics. In: Stephanidis C, Antona M, editors. Universal access in human-computer interaction. Design for all and accessibility practice. Cham, Switzerland: Springer International Publishing; 2014:613-622.
86. Rapp A, Hopfgartner F, Plumbaum T, Kay J, Kummerfeld B, Herder E. LinkQS 2014: Linking the Quantified Self Workshop. In: Hypertext 2014: Extended Proceedings of the 25th ACM Hypertext and Social Media Conference.: CEUR-WS; 2014 Presented at: Hypertext 2014; September 1 - 4, 2014; Santiago, Chile.
87. Reigeluth T. Why data is not enough: Digital traces as control of self and self-control. Surveillance & Society. 2014. (2) p. 243-254 URL: http://ojs.library.queensu.ca/index.php/surveillance-and-society/article/view/enough [WebCite Cache ID 6hd1oetJ0]
88. Salamati F, Pasek Z. Personal wellness: Complex and elusive product and distributed self-services. Procedia CIRP 2014;16:283-288. [doi: 10.1016/j.procir.2014.02.016]
89. Schreier G. The internet of things for personalized health. Stud Health Technol Inform 2014;200:22-31. [Medline: 24851958]
90. Shull P, Jirattigalachote W, Hunt M, Cutkosky M, Delp S. Quantified self and human movement: a review on the clinical impact of wearable sensing and feedback for gait analysis and intervention. Gait Posture 2014;40(1):11-19. [doi: 10.1016/j.gaitpost.2014.03.189] [Medline: 24768525]

91. Sjöklint M. The measurable me: the influence of self-quantification on the online user's decision-making process. In: Proceedings of the 2014 ACM International Symposium on Wearable Computers: Adjunct Program ISWC.: ACM; 2014 Presented at: UbiComp '14 - 2014 ACM Conference on Ubiquitous Computing, Adjunct Program; September 13 - 17, 2014; Seattle, WA p. 131-137. [doi: 10.1145/2641248.2642737]
92. Wenger M, Bell J, McEvoy P, Yamaguchi C, Shokrpour A. Bloom: fostering healthy and peaceful pregnancies with personal analytics. In: CHI EA '14 Extended Abstracts on Human Factors in Computing Systems.: ACM; 2014 Presented at: CHI '14 ACM Conference on Human Factors in Computing Systems; April 26, - May 1, 2014; Toronto, Ontario p. 245-250. [doi: 10.1145/2559206.2580935]
93. Zhu Z, Blanke U, Calatroni A, Brdiczka O, Tröster G. Fusing on-body sensing with local and temporal cues for daily activity recognition. In: Proceedings of the 9th International Conference on Body Area Networks. 2014 Presented at: BODYNETS '14: 9th International Conference on Body Area Networks; September 29 - October 1, 2014; London p. 83-89. [doi: 10.4108/icst.bodynets.2014.257014]
94. Altini M, Penders J, Amft O. Body weight-normalized energy expenditure estimation using combined activity and allometric scaling clustering. In: Conf Proc IEEE Eng Med Biol Soc.: IEEE; 2013 Presented at: Proceedings of the 35th Annual International Conference of IEEE Engineering in Medicine and Biology Society, EMBC; July 3 - 7, 2013; Osaka p. 6752-6755. [doi: 10.1109/EMBC.2013.6611106]
95. Buzzo D. Lost time never. In: Proceedings of the 2013 Inputs-Outputs Conference: An Interdisciplinary Conference on Engagement in HCI and Performance.: ACM; 2013 Presented at: Inputs-Outputs '13 - 2013 Inputs-Outputs Conference: An Interdisciplinary Conference on Engagement in HCI and Performance; June 26, 2013; Brighton, United Kingdom p. 1-4. [doi: 10.1145/2557595.2557604]
96. Calvo R, Peters D. The irony and re-interpretation of our quantified self. In: Proceedings of the 25th Australian Computer-Human Interaction Conference: Augmentation, Application, Innovation, Collaboration.: ACM; 2013 Presented at: OzCHI '13 - 25th Australian Computer-Human Interaction Conference; November 25 - 29, 2013; Adelaide p. 367-370. [doi: 10.1145/2541016.2541070]
97. Chen Z, Lin M, Chen F, Lane N, Cardone G, Wang R. Unobtrusive sleep monitoring using smartphones. In: Proceedings of the 2013 7th International Conference on Pervasive Computing Technologies for Healthcare and Workshops.: IEEE; 2013 Presented at: 7th International Conference on Pervasive Computing Technologies for Healthcare and Workshops; May 5 - 8, 2013; Venice p. 145-152.
98. Cuttone A, Lehmann S, Larsen J. A mobile personal informatics system with interactive visualizations of mobility and social interactions. In: Proceedings of the 1st ACM international workshop on Personal data meets distributed multimedia.: ACM; 2013 Presented at: PDM '13 - 1st ACM international workshop on Personal data meets distributed multimedia; October 21 - 25, 2013; Barcelona p. 27-30. [doi: 10.1145/2509352.97]
99. Derksen G, Ruecker S, Causer T, Terras M. Demonstrating data using storyboard visualization tool. In: Proceedings of the 6th International Symposium on Visual Information Communication and Interaction.: ACM; 2013 Presented at: VINCI '13 - 6th International Symposium on Visual Information Communication and Interaction; August 17 - 18, 2013; Tianjin, China p. 117-117. [doi: 10.1145/2493102.2493115]
100. Jain R. What's in it for me? how can big multimedia aid quantified-self applications. In: Proceedings of the 1st ACM international workshop on Personal data meets distributed multimedia.: ACM; 2013 Presented at: PDM '13 - 1st ACM international workshop on Personal data meets distributed multimedia; October 21 - 25, 2013; Barcelona p. 43-44. [doi: 10.1145/2509352.2509403]
101. Kido T, Swan M. Exploring the mind with the aid of personal genome: citizen science genetics to promote positive well-being. : AAAI; 2013 Presented at: Data Driven Wellness - AAAI Spring Symposium Series; 2013; Palo Alto, CA URL: http://www.aaai.org/ocs/index.php/SSS/SSS13/paper/view/5736 [WebCite Cache ID 6hd1vTfwf]
102. Kunze K, Iwamura M, Kise K, Uchida S, Omachi S. Activity recognition for the mind: Toward a cognitive "Quantified Self". Computer 2013 Oct;46(10):105-108. [doi: 10.1109/MC.2013.339]
103. Lupton D. Quantifying the body: monitoring and measuring health in the age of mHealth technologies. Critical Public Health 2013 Dec;23(4):393-403. [doi: 10.1080/09581596.2013.794931]
104. Lupton D. Understanding the Human Machine [Commentary]. IEEE Technol. Soc. Mag 2013 Dec;32(4):25-30. [doi: 10.1109/mts.2013.2286431]
105. Martin-Sanchez F, Lopez-Campos G, Gray K. Biomedical informatics methods for personalized medicine and participatory health. In: Sarkar IN, editor. Methods in biomedical informatics: a pragmatic approach. Amsterdam: Elsevier/AP; 2014:347-394.
106. Matassa A, Rapp A, Simeoni R. Wearable accessories for cycling: tracking memories in urban spaces. In: UbiComp Adjunct '13 Proceedings of the 2013 ACM Conference on Pervasive and Ubiquitous Computing adjunct publication. New York: ACM; 2013 Presented at: ACM Conference on Pervasive and Ubiquitous Computing; September 8-12, 2013; Zurich, Switzerland p. 415-424. [doi: 10.1145/2494091.2495973]
107. Mcfedries P. Tracking the quantified self [Technically speaking]. IEEE Spectr 2013 Aug;50(8):24. [doi: 10.1109/mspec.2013.6565555]

108. Pasek Z, Bacioiu G. Self-service for personal health monitoring and decisions. In: Spohrer JC, Freund LE, editors. Advances in the Human Side of Service Engineering. Boca Raton, FL: CRC Press; 2012:13-20.
109. Salamati F, Pasek Z. Modeling for personal well-being: time for paradigm change. In: Proceedings of the 2013 Grand Challenges on Modeling and Simulation Conference.: Society for Modeling & Simulation International; 2013 Presented at: GCMS '13 - Grand Challenges on Modeling and Simulation Conference; July 7 - 10, 2013; Toronto, Ontario p. 11.
110. Swan M. Next-generation personal genomic studies: extending social intelligence genomics to cognitive performance genomics in quantified creativity and thinking fast and slow. : AAAI; 2013 Presented at: Data Driven Wellness - AAAI Spring Symposium Series; 2013; Palo Alto, CA.
111. Yumak Z, Pu P. Survey of Sensor-Based Personal Wellness Management Systems. BioNanoSci 2013 Jul 10;3(3):254-269. [doi: 10.1007/s12668-013-0099-0]
112. Bottles K. Will the quantified self movement take off in health care? Physician Exec 2012;38(5):74-75. [Medline: 23888667]
113. Lathia N. Using idle moments to record your health via mobile applications. In: Proceedings of the 1st ACM workshop on Mobile systems for computational social science.: ACM; 2012 Presented at: 1st ACM workshop on Mobile systems for computational social science; June 25, 2012; Low Wood Bay, Lake District, UK p. 22-27. [doi: 10.1145/2307863.2307871]
114. Lathia N. Using ratings to profile your health. In: Proceedings of the 6th ACM conference on Recommender systems.: ACM; 2012 Presented at: RecSys '12 - 6th ACM conference on Recommender systems; September 9 - 13, 2013; Dublin p. 303-304. [doi: 10.1145/2365952.2366024]
115. Li I, Medynskiy Y, Froehlich J, Larsen J. Personal informatics in practice: improving quality of life through data. In: CHI '12 Extended Abstracts on Human Factors in Computing Systems.: ACM; 2012 Presented at: CHI '12 - CHI Conference on Human Factors in Computing Systems; May 5 - 10, 2012; Austin, TX p. 2799-2802. [doi: 10.1145/2212776.2212724]
116. Smarr L. Quantifying your body: a how-to guide from a systems biology perspective. Biotechnol J 2012 Aug;7(8):980-991. [doi: 10.1002/biot.201100495] [Medline: 22887886]
117. Rivera-Pelayo V, Zacharias V, Mller L, Braun S. Applying quantified self approaches to support reflective learning. In: Proceedings of the 2nd International Conference on Learning Analytics and Knowledge.: ACM; 2012 Presented at: LAK '12 - 2nd International Conference on Learning Analytics and Knowledge; April 29 - May 2, 2012; Vancouver p. 111-114. [doi: 10.1145/2330601.2330631]
118. Swan M. Crowdsourced health research studies: an important emerging complement to clinical trials in the public health research ecosystem. J Med Internet Res 2012 Mar;14(2):e46 [FREE Full text] [doi: 10.2196/jmir.1988] [Medline: 22397809]
119. Swan M. Health 2050: the realization of personalized medicine through crowdsourcing, the quantified self, and the participatory biocitizen. Journal of Personalized Medicine 2012;2(3):93-118. [doi: 10.3390/jpm2030093] [Medline: 25562203]
120. Swan M. Sensor Mania! The Internet of Things, Wearable Computing, Objective Metrics, and the Quantified Self 2.0. JSAN 2012 Nov 08;1(3):217-253. [doi: 10.3390/jsan1030217]
121. The Economist. 2012 Mar 03. The quantified self: counting every moment URL: http://www.economist.com/node/21548493 [accessed 2015-07-12] [WebCite Cache ID 6ZxEeccdf]
122. Swan M. Emerging patient-driven health care models: an examination of health social networks, consumer personalized medicine and quantified self-tracking. Int J Environ Res Public Health 2009 Feb;6(2):492-525 [FREE Full text] [doi: 10.3390/ijerph6020492] [Medline: 19440396]
123. Thomas J, Harden A. Methods for the thematic synthesis of qualitative research in systematic reviews. BMC Med Res Methodol 2008;8:45 [FREE Full text] [doi: 10.1186/1471-2288-8-45] [Medline: 18616818]
124. 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]
125. Elo S, Kaariainen M, Kanste O, Polkki T, Utriainen K, Kyngas H. Qualitative Content Analysis: A Focus on Trustworthiness. SAGE Open 2014 Feb 11;4(1):1-10. [doi: 10.1177/2158244014522633]
126. Kaptelinin V. The Object of activity: Making sense of the sense-maker. Mind, Culture, and Activity 2005 Feb;12(1):4-18. [doi: 10.1207/s15327884mca1201_2]
127. Liang Z, Martell M. Framing self-quantification for individual-level preventive health care. In: Proceedings of the 8th International Conference on Health Informatics HEALTHINF 2015. Setúbal, Portugal: SCITEPRESS Digital Library; 2015 Presented at: 8th International Conference on Health Informatics HEALTHINF 2015; January 12-15, 2015; Lisbon p. 50. [doi: 10.5220/0005202503360343]
128. Oh J, Lee U. Exploring UX issues in quantified self technologies. In: Eighth International Conference on Mobile Computing and Ubiquitous Networking (ICMU).: IEEE; 2015 Presented at: 8th International Conference on Mobile Computing and Ubiquitous Networking, ICMU; 20-22 Jan, 2015; Hakodate p. 53-59. [doi: 10.1109/ICMU.2015.7061028]
129. Whooley M, Ploderer B, Gray K. On the integration of self-tracking data amongst quantified self members. In: Proceedings of the 28th International BCS Human Computer Interaction Conference HCI 2014; Southport, UK.: BCS; 2014 Presented at: 28th International BCS Human Computer Interaction Conference HCI 2014; 9-12 September 2014; Southport, UK p. 151-160. [doi: 10.14236/ewic/hci2014.16]
130. Choe EK, Lee B, Schraefel MC. Characterizing Visualization Insights from Quantified-Selfers' Personal Data Presentations. IEEE Comput Graph Appl 2015 May 13;35(4):28-37. [doi: 10.1109/MCG.2015.51]

131. Kim J. Analysis of health consumers' behavior using self-tracker for activity, sleep, and diet. Telemed J E Health 2014 Jun;20(6):552-558 [FREE Full text] [doi: 10.1089/tmj.2013.0282] [Medline: 24745608]
132. Doyle J, Walsh L, Sassu A, McDonagh T. Designing a wellness self-management tool for older adults: results from a field trial of YourWellness. In: PervasiveHealth '14 Proceedings of the 8th International Conference on Pervasive Computing Technologies for Healthcare.: ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering); 2014 Presented at: 8th International Conference on Pervasive Computing Technologies for Healthcare; May 20-23, 2014; Oldenberg, Germany p. 134-141. [doi: 10.4108/icst.pervasivehealth.2014.254950]
133. Papi E, Belsi A, McGregor AH. A knee monitoring device and the preferences of patients living with osteoarthritis: a qualitative study. BMJ Open 2015 Sep;5(9):e007980 [FREE Full text] [doi: 10.1136/bmjopen-2015-007980] [Medline: 26346873]
134. Shih P, Han K, Poole E, Rosson M, Carroll J. iConference 2015 Proceedings.: iSchools; 2015 Mar 15. Use and adoption challenges of wearable activity trackers URL: http://hdl.handle.net/2142/73649 [WebCite Cache ID 6hTLKget4]
135. Ivanov A, Sharman R, Rao HR. Exploring factors impacting sharing health-tracking records. Health Policy and Technology 2015 Sep;4(3):263-276. [doi: 10.1016/j.hlpt.2015.04.008]
136. Chang J. Self-tracking for distinguishing evidence-based protocols in optimizing human performance and treating chronic illness. In: Self tracking and collective intelligence for personal wellness: Papers from the 2012 AAAI Spring Symposium. Palo Alto, CA: AAAI Press; 2012 Presented at: AAAI 2012 Spring Symposium: Self-Tracking and Collective Intelligence for Personal Wellness; 2012; Palo Alto, CA.
137. Lee V. What's happening in the "Quantified Self" movement? In: Learning and becoming in practice: proceedings of International Conference of the Learning Sciences, ICLS 2014.: International Society of the Learning Sciences; 2014 Presented at: 11th International Conference of the Learning Sciences, ICLS 2014; June 23-27, 2014; Boulder, CO p. 1032-1036 URL: http://digitalcommons.usu.edu/itls_facpub/491 [WebCite Cache ID 6hd208Bmx]
138. Packer H, Buzogany G, Smith D, Dragan L, Kleek M, Shadbolt N. The editable self: a workbench for personal activity data. In: CHI EA '14 Extended Abstracts on Human Factors in Computing Systems. New York: ACM; 2014 Presented at: CHI '14; April 26 - May 1, 2014; Toronto p. 2185-2190. [doi: 10.1145/2559206.2581283]
139. Guo F, Li Y, Kankanhalli M, Brown M. An evaluation of wearable activity monitoring devices. In: Proceedings of the 1st ACM international workshop on Personal data meets distributed multimedia; Barcelona, Spain. New York: ACM; 2013 Presented at: PDM '13 Proceedings of the 1st ACM international workshop on Personal data meets distributed multimedia; 21-25 October 2013; Barcelona, Spain p. 31-34. [doi: 10.1145/2509352.2512882]
140. Punnoose B, Gray K. Comparative evaluation of two systems for integrating biometric data from self-quantification. In: Health Information Science. Lecture Notes in Computer Science. New York: Springer International Publishing; 2015:195-201.
141. Lee V, Drake J. Quantified recess: design of an activity for elementary students involving analyses of their own movement data. In: IDC '13: Proceedings of the 12th International Conference on Interaction Design and Children. New York: ACM; 2013 Presented at: 12th International Conference on Interaction Design and Children; 24-27 June 2013; New York p. 273-276. [doi: 10.1145/2485760.2485822]
142. Dontje ML, de GM, Lengton RR, van der Schans Cees P, Krijnen WP. Measuring steps with the Fitbit activity tracker: an inter-device reliability study. J Med Eng Technol 2015 Jul;39(5):286-290. [doi: 10.3109/03091902.2015.1050125] [Medline: 26017748]
143. Pickard K, Swan M. Big desire to share big health data: A shift in consumer attitudes toward personal health information. In: Big data becomes personal: Knowledge into meaning - 2014 AAAI Spring Symposium Series. Palo Alto, CA: AAAI Press; 2014 Presented at: 2014 AAAI Spring Symposium; 2014; Palo Alto, CA p. 51-58. [doi: 10.13140/2.1.2107.2960]
144. Jones W. Personal Information Management. Ann Rev Info Sci Tech 2008 Oct 24;41(1):453-504. [doi: 10.1002/aris.2007.1440410117]
145. Hibbard JH, Mahoney ER, Stock R, Tusler M. Do increases in patient activation result in improved self-management behaviors? Health Serv Res 2007 Aug;42(4):1443-1463 [FREE Full text] [doi: 10.1111/j.1475-6773.2006.00669.x] [Medline: 17610432]
146. Hibbard JH, Peters E, Dixon A, Tusler M. Consumer competencies and the use of comparative quality information: it isn't just about literacy. Med Care Res Rev 2007 Aug;64(4):379-394. [doi: 10.1177/1077558707301630] [Medline: 17684108]
147. Fowles JB, Terry P, Xi M, Hibbard J, Bloom CT, Harvey L. Measuring self-management of patients' and employees' health: further validation of the Patient Activation Measure (PAM) based on its relation to employee characteristics. Patient Educ Couns 2009 Oct;77(1):116-122. [doi: 10.1016/j.pec.2009.02.018] [Medline: 19356881]
148. Mursu A, Luukkonen I, Toivanen M, Korpela M. Activity theory in information systems research and practice: theoretical underpinnings for an information systems development model. Information Research. 2006. (3) p. 3 URL: http://www.informationr.net/ir/12-3/paper311.html [WebCite Cache ID 6hd2A59EY]
149. Kaptelinin V, Nardi B. Activity theory: basic concepts and applications. In: CHI EA '97 Extended Abstracts on Human Factors in Computing Systems. New York: ACM; 1997 Presented at: CHI '97 Human Factors in Computing Systems; 1997; Atlanta, Georgia p. 158-159. [doi: 10.1145/1120212.1120321]
150. Bedny G, Karwowski W. Activity theory as a basis for the study of work. Ergonomics 2004 Feb 5;47(2):134-153. [doi: 10.1080/00140130310001617921] [Medline: 14660209]

151. Eurostat. 2014. Measuring digital skills across the EU: EU wide indicators of digital competence URL: http://ec.europa.eu/information_society/newsroom/cf/document.cfm?action=display&doc_id=5406 [accessed 2016-05-13] [WebCite Cache ID 6hd2YA1ly]
152. Kayser L, Kushniruk A, Osborne RH, Norgaard O, Turner P. Enhancing the effectiveness of consumer-focused health information technology systems through eHealth literacy: a framework for understanding users' needs. JMIR Human Factors 2015 May 20;2(1):e9. [doi: 10.2196/humanfactors.3696] [Medline: 27025228]
153. Sherer S. Patients are not simply health it users or consumers: the case for “e Healthicant” applications. Communications of the Association for Information Systems 2014;34(1):351-364.
154. Almalki M, Sanchez F, Gray K. Quantifying the Activities of Self-quantifiers: Management of Data, Time and Health. Stud Health Technol Inform 2015;216:333-337. [Medline: 26262066]
155. Almalki M, Gray K, Martin-Sanchez F. Classification of data and activities in self-quantification systems. 2014 Presented at: HISA Big Data in Biomedicine and Healthcare Conference; April 3-4, 2014; Melbourne URL: http://ceur-ws.org/Vol-1149/bd2014_almalki.pdf [WebCite Cache ID 6hd2EraNH]

Abbreviations

AT: Activity Theory
DM: data management
HM: health management
ICT: information and communication technology
SQ: self-quantification

Edited by G Eysenbach; submitted 03.08.15; peer-reviewed by S Smith, D Vollmer Dahlke, S Kelders; comments to author 07.09.15; revised version received 16.11.15; accepted 21.03.16; published 27.05.16.

Please cite as:
Almalki M, Gray K, Martin-Sanchez F
Activity Theory as a Theoretical Framework for Health Self-Quantification: A Systematic Review of Empirical Studies
J Med Internet Res 2016;18(5):e131
URL: http://www.jmir.org/2016/5/e131/
doi:10.2196/jmir.5000
PMID:27234343

Corrigenda and Addenda

Correction: Online Alcohol Assessment and Feedback for Hazardous and Harmful Drinkers: Findings From the AMADEUS-2 Randomized Controlled Trial of Routine Practice in Swedish Universities

Preben Bendtsen[1*], PhD; Marcus Bendtsen[2*], MSc (Med Info); Nadine Karlsson[1*], PhD; Ian R White[3*], PhD; Jim McCambridge[4], PhD

[1]Department of Medical and Health Sciences, Linköping University, Linköping, Sweden

[2]Department of Computer Science, University of Linköping, Linköping, Sweden

[3]Cambridge Institute of Public Health, MRC Biostatistics Unit, Cambridge, United Kingdom

[4]Faculty of Science, Department of health Sciences, University of York, Heslington, United Kingdom

[*]these authors contributed equally

Corresponding Author:
Preben Bendtsen, PhD
Department of Medical and Health Sciences
Linköping University
Campus US
Linköping, 581 83
Sweden
Phone: 46 702324615
Fax: 46 702324615
Email: preben.bendtsen@liu.se

Related Article:

Correction of: http://www.jmir.org/2015/7/e170/

(J Med Internet Res 2016;18(5):e118) doi:10.2196/jmir.5724

The authors of the paper “Online Alcohol Assessment and Feedback for Hazardous and Harmful Drinkers: Findings From the AMADEUS-2 Randomized Controlled Trial of Routine Practice in Swedish Universities” (J Med Internet Res 2015;17(7):e170) used the wrong variable for the number of email reminders before participants answered the follow-up. In Table 2, the values in columns two, three, and four under the subcategory "Number of follow-up emails before response, n (%)" should read as in the table below. A footnote has also been added to the *P*-value associated with this variable. In the Results section, there remains no statistically significant association between the primary outcome and the number of email reminders (*P*=.16), consistent with the missing at random (MAR) assumption. However, analyses using the repeated attempts model and linear regression suggested an intervention effect of an 11% reduction (18% reduction to 3% reduction) and a statistically significant departure from the MAR assumption. The authors are unable to explain why this result contradicts the above analysis that uses the same variables, so they prefer to trust the simpler model and hence they do not regard the repeated attempts model findings as providing any strong evidence against the MAR assumption. This error has been corrected in the online version of the paper on the JMIR website on May 16, 2016 together with publishing this correction notice. A correction notice has been sent to PubMed and PubMed Central.

Table 2. Comparison of groups at follow-up (total n=931 without University of Gävle).

Number of follow-up emails before response, n (%)	Intervention (n=402)	Control (n=529)	*P*-value
1	222 (55.2)	315 (59.6)	.07 [a]
2	82 (20.4)	103 (19.5)	
3	51 (12.7)	70 (13.2)	
4	27 (6.7)	23 (4.3)	
5	20 (5.0)	18 (3.4)	

[a]Trend test.

Edited by G Eysenbach; submitted 06.03.16; this is a non–peer-reviewed article;accepted 08.03.16; published 16.05.16.

Please cite as:
Bendtsen P, Bendtsen M, Karlsson N, White IR, McCambridge J
Correction: Online Alcohol Assessment and Feedback for Hazardous and Harmful Drinkers: Findings From the AMADEUS-2 Randomized Controlled Trial of Routine Practice in Swedish Universities
J Med Internet Res 2016;18(5):e118
URL: http://www.jmir.org/2016/5/e118/
doi:10.2196/jmir.5724
PMID:27183182

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org